Open Journal of Statistics, 2017, 7, 122-131
http://www.scirp.org/journal/ojs
ISSN Online: 2161-7198
ISSN Print: 2161-718X

Forecasting Foreign Direct Investment to Zambia: A Time Series Analysis

Stanley Jere*, Bornwell Kasense, Obvious Chilyabanyama

Department of Mathematics and Statistics, Mulungushi University, Kabwe, Zambia
Email: *sjere@mu.ac.zm

How to cite this paper: Jere, S., Kasense, B. and Chilyabanyama, O. (2017) Forecasting Foreign Direct Investment to Zambia: A Time Series Analysis. *Open Journal of Statistics*, **7**, 122-131.
https://doi.org/10.4236/ojs.2017.71010

Received: October 17, 2016
Accepted: February 25, 2017
Published: February 28, 2017

Abstract

Three methods are considered in this paper: Simple exponential smoothing (SES), Holt-Winters exponential smoothing (HWES) and autoregressive integrated moving average (ARIMA). The best fit model was then used to forecast Zambia's annual net foreign direct investment (FDI) inflows from 1970 to 2014. Foreign direct investment is foreign capital investment to Zambia. Throughout the paper the methods are illustrated using Zambia's annual Net FDI inflows. A comparison of the three methods shows that the ARIMA (1, 1, 5) is the best fit model because it has the minimum error. Forecasting results give a gradual increase in annual net FDI inflows of about 44.36% by 2024. Forecasting results plays a vital role to policy makers. Decision making, coming up with good policies and suitable strategic plans, depends on accurate forecasts. Zambian FDI policy makers can use the results obtained in this study and create suitable strategic plans to promote FDI.

Keywords

Foreign Direct Investment, Simple Exponential Smoothing, Holt-Winters Exponential Smoothing, Autoregressive Integrated Moving Average, Forecasting

1. Introduction

Forecasting is key in many fields of science. In this paper, three methods are of interest and they are: SES, HWES and ARIMA. Throughout the paper the methods are illustrated using Zambia's annual net FDI inflows. The best fit model is used to forecast Zambia's annual net foreign direct investment (FDI) from 1970 to 2014.

SES method is a simple tool for forecasting time series data. Smoothing implies removing the unwanted noise so that the general path is created. This method is suitable for forecasting data without trend or seasonal pattern. It is basically a recursive computing procedure [1]. The HWES method prediction items are obtained as a weighted average of past observed values where the weights reduce exponentially so that the values of recent observations contribute to the forecast more than the values of earlier observations [2]. ARIMA models can be used to produce forecasts for time series data. The ARIMA model has three parts. Not all parts are always necessary but it depends on the type of time series data at hand. The three parts are the autoregressive (AR), the integrated (I) and lastly, the moving average (MA). Assumption for the AR part of a time series data is that the observed value depends on some linear combinations of previous observed values up to some maximum lags plus error term. Assumption for the MA part of time series data is that the observed value is a random error term plus some linear combinations of previous random error terms up to some maximum lags [3].

FDI is foreign capital investment to a country. Foreign direct investment to a country results in increasing productivity, reducing unemployment, and increasing the use of technology. The need for FDI came as a result of shortages in domestic funding sources to finance development projects in developing countries. These developing countries realized that it is through FDI that they can achieve economic growth. According to [4], FDI excludes loan from international organizations, foreign governments, private commercial banks, stocks and bonds purchased by foreigners but it is an investment where managerial control is done by foreign investors. According to [5] FDI constitute activities such as decision making that are done by firms or groups of firms outside to the country of investment. Furthermore [6], FDI is defined as an investment that arises when the investor in the mother country invests in another country with an intension to have control on how to manage and run it.

Studies by [7] indicate that there is a positive relationship between FDI and economic growth. However, this positive relationship depends on human capital available in that economy. Furthermore, for countries with very low levels of human capital, then the direct effect of FDI is negative. They further argued that FDI result to competition with domestic investors and as a result the local and already existing businesses are affected negatively hence the weak positive relationship with Economic Growth. Studies by [8] indicate that, FDI inflows in developing countries led to "crowd in" other investment at macro level. Studies by [9] found that FDI inflows led to higher per capita GDP, economic growth rate, productivity growth, higher export in host country and increased backward, forward linkages with affiliates to multinationals.

The main goal for the Zambia government is to increase and sustain FDI inflows beyond the current levels to highly benefit the country. Zambian FDI inflows are mainly in copper and cobalt extraction, agricultural sector particular in horticulture and floriculture production, and in tourism. Firms or groups of firms from countries like United Kingdom and South Africa have traditionally been the main contributors of foreign direct investment though FDI inflow from other countries drastically increases. The net inflow to other countries is negative indicating outflows that are FDI from those countries (inflows) are less than those from Zambia (outflows). However, the scope of this research is to discuss FDI inflows (Net) to Zambia. An analysis of the FDI flows by source country in 2012 shows that Canada (US $724.3 million), South Africa (US $426.0 million), the Netherlands (US $262.2 million) and the United Kingdom (US $227.2 million), were the major source countries of Zambia's FDI inflows, accounting for 94.7 per cent of total inflows, collectively. The other source countries are Switzerland (US $166.9 million), China (US $141.9 million), Nigeria (US $94.6 million), Singapore (US $62.0 million), Congo DR (US $28.6 million) and France (US $20.2 million) [10] [11] [12].

Forecasting results plays a vital role to policy makers. Decision making, coming up with good policies and suitable strategic plans, depends on accurate forecasts [13].

2. Methodology

2.1. Simple Exponential Smoothing Model (SES)

Simple exponential smoothing method involves smoothing out random fluctuations of time series data. The method is suitable for forecasting data without trend or seasonal pattern. This method gives past data weights known as smoothing constants that decrease exponentially with time. Below is the exponential smoothing model for time series data X_t is shown below:

$$\bar{X}_t = \alpha X_t + (1-\alpha)\bar{X}_{t-1} \tag{1}$$

where α *is the smoothing constant,* $0<\alpha<1$, $t=1,2,\cdots,T$, X_t *is raw time series data and* $\bar{X}_t$ *is smoothed data or output.*

The h *-step-ahead forecast equation is*

$$\hat{X}_{t+h} = \bar{X}_t \tag{2}$$

where $h=1,2,3,\cdots$ [13].

2.2. Holt-Winters Exponential Smoothing Model (HWES)

Holt-Winters exponential smoothing method is an extension of SES and uses a linear combination of the previous values of a series for generating and modeling future values. It is applies to time series data that has trend. Recent time series recordings are key to forecasting future values of a series. The model for time series data X_t is as shown below:

$$\begin{aligned}&\bar{X}_t = \alpha X_t + (1-\alpha)\left(\bar{X}_{t-1} + b_{t-1}\right)\\&0<\alpha<1\end{aligned} \tag{3}$$

$$\begin{aligned}&b_t = \beta\left(\bar{X}_t - \bar{X}_{t-1}\right) + (1-\beta)b_{t-1}\\&0<\beta<1\end{aligned} \tag{4}$$

where α is the smoothing constant, β is the trend smoothing constants, X_t is raw data, $\bar{X}_t$ is smoothed data and b_t is the trend estimates.

The h -step-ahead forecast equation is

$$\widehat{X}_{t+h} = \overline{X}_t + hb_t \quad ([13]) \tag{5}$$

2.3. Autoregressive Integrated Moving Average Model (ARIMA)

Stochastic models attributed to Box-Jenkins known as the ARIMA have been found to be more efficient and reliable even for short term forecasting. Further, stochastic models are distribution-free as no assumptions are required about the data [14]. The ARIMA model consists of the following expressions called the order of autoregressive (AR) model (*p*), differencing order (*d*) and the order of moving average (MA) model (*q*). The Box-Jenkin models are denoted by ARIMA (*p*, *d*, *q*). "*I*" implies that the process need to undergo differentiation and when the modelling is done, the results undergo an integration process to produce forecasts and estimates. The expressions for MA, AR and ARMA are as follows:

$$\text{AR model: } \widehat{X}_t = \sum_{i=1}^{p} \vartheta_i X_{t-i} + \varepsilon_t \tag{6}$$

$$\text{MA model: } \widehat{X}_t = \sum_{i=1}^{q} \varphi_i \varepsilon_{t-i}, \tag{7}$$

$$\text{ARMA model: } \widehat{X}_t = \sum_{i=1}^{p} \vartheta_i \mathrm{X}_{t-i} + \varepsilon_t + \sum_{i=1}^{q} \varphi_i \varepsilon_{t-i} \tag{8}$$

where ϑ_t is the autoregressive parameter at time *t*, ε_t is the error term at time *t* and φ_t is the moving-average parameter at time *t* [13].

2.4. The Error Measures for Model-Selection

The error measures are used to compare how well models fit the time series. According to [13], the best fit or forecasting model is one with minimal errors. The following error indicators were used in this paper:

3. Results and Discussion

The SES, HWES and ARIMA models are used to forecasting Zambia's annual Net Foreign direct investment (FDI) inflows from 1970 to 2014. *R* is a widely used statistical software package for statistical analysis. It was used to come up with SES, HWES and ARIMA models. *R* contains built-in functions that allow the user to determine model parameters spontaneously; the only requirement in this software is the time series data to be analysed. Using *R*, the SES model indicates that the parameter $\alpha = 0.73$ is the best parameter value. The equation for this model thus takes the form

$$\overline{X}_t = 0.73X_t + 0.27\overline{X}_{t-1} \tag{9}$$

The HWES model indicate that the parameters $\alpha = 0.31$ and $\beta = 0.41$, giving us the following equations:

$$\overline{X}_t = 0.31X_t + 0.69\left(\overline{X}_{t-1} + b_{t-1}\right) \tag{10}$$

$$b_t = 0.41\left(\overline{X}_t - \overline{X}_{t-1}\right) + 0.59b_{t-1} \tag{11}$$

For ARIMA model, the procedure is achieved by considering the following steps: identification, model selection, parameter estimation and diagnostic check [13]. The steps are illustrated below:

Step 1: ARIMA model identification:

Time plot is the first step of ARIMA model identification of time series. A time plot of the FDI is plotted in **Figure 1** for $d = 0$ and $d = 1$. Stationarity can now be checked using visual display of the ACF and PACF graphs in **Figure 2**. The ACF and PACF plots in **Figure 2** show that the FDI's time series data are not stable for $d = 0$ due to its slow decay and therefore nonstationary. For $d = 1$, the time plot is stationary. According to [15] [16], converting a nonstationary time series to a stationary one through differencing (where needed) is an important part of the process of fitting an ARIMA model.

Step 2: Model selection

The ACF and PACF plots for $d = 1$ in **Figure 2** indicate that the first differenced FDI series are stationary hence require further examination to establish the most suitable ARIMA. **Table 1** shows the formula for each error indicator considered in this study. **Table 2** shows the details of various ARIMA models along the error measures. Results by [14] demonstrate that an ARIMA model with lowest error measures specifically the AIC is considered the best model for forecasting. In this case an ARIMA (1, 1, 5) is considered as best fit model because it has the lowest value of the AIC statistics.

Step 3: Model fitting and Parameter estimation.

R output (version 0.99.903) for estimated parameter and *p*-value:

Arima (1, 1, 5))

Coefficients:

ar1 ma1 ma2 ma3 ma4 ma5

0.8300 −1.3240 0.4039 0.5289 −0.8271 0.6342

sigma^2 estimated as 27972: log likelihood = −290.67, aic = 595.33

The parameters found to be significance at 5% in **Table 3** are AR (1), MA (1), MA (2), MA (3), MA (4), and MA (5). The ARIMA (1, 1, 5) model equation can

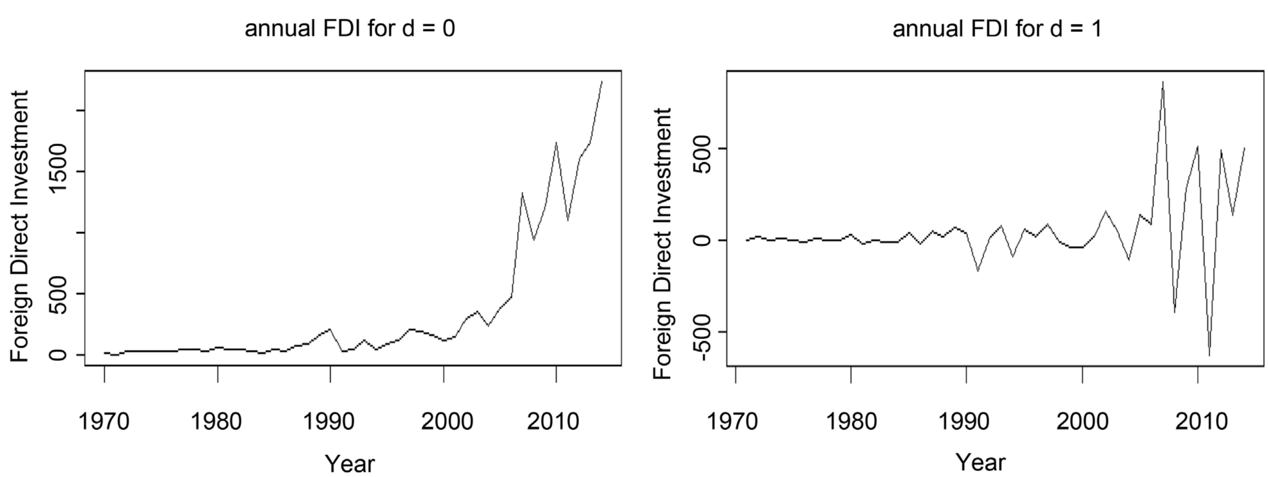

Figure 1. Time plots for $d = 0$ and $d = 1$.

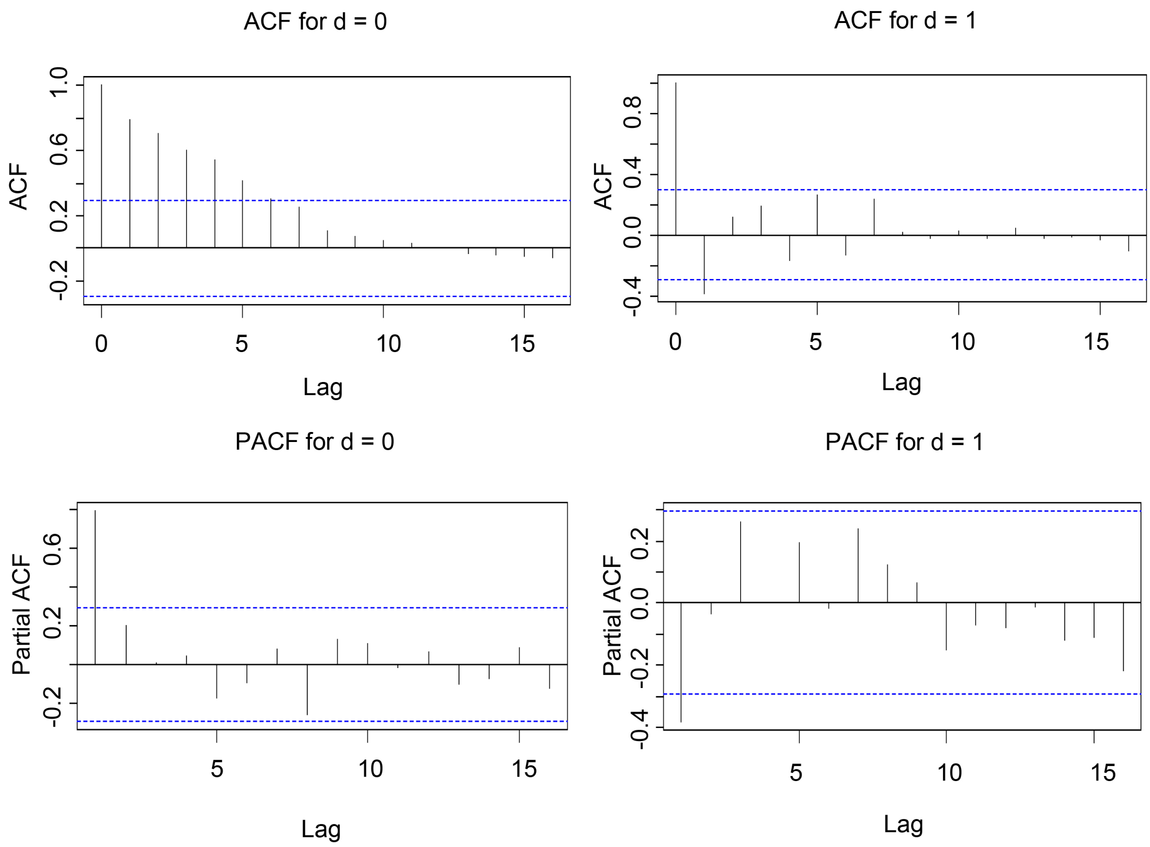

Figure 2. Plots of ACF and PACF for $d = 0$ and $d = 1$.

Table 1. The error indicators.

Criteria	Formula	Criteria	Formula
MPE	$\frac{1}{n}\sum_{i=1}^{n}100\times\frac{\varepsilon_i}{y_i}$	RMSE	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i^2}$
MAE	$\frac{1}{n}\sum_{i=1}^{n}\lvert\varepsilon_i\rvert$	MAPE	$\frac{1}{n}\sum_{i=1}^{n}\left\lvert\frac{\varepsilon_i}{x_i}\right\rvert\times 100$
MASE	$\left\lvert\frac{\varepsilon_i}{\frac{1}{n-1}\sum_{i=1}^{n}\lvert Y_t - Y_{i-1}\rvert}\right\rvert$		

Table 2. The AIC statistics measures for selected ARIMA models.

TENTATIVE MODEL	**ARIMA (1,1,5)**	**ARIMA (5,1,1)**	**ARIMA (1,1,3)**	**ARIMA (5,1,0)**	**ARIMA (0,1,2)**	**ARIMA (3,1,0)**	**ARIMA (1,1,0)**	**ARIMA (0,1,5)**
AIC	**595.33**	598.89	599.69	599.76	600.25	600.29	600.69	600.74
TENTATIVE MODEL	ARIMA (0,1,3)	ARIMA (4,1,0)	ARIMA (0,1,1)	ARIMA (0,1,4)	ARIMA (2,1,0)	ARIMA (1,1,1)	ARIMA (2,1,2)	
AIC	601.55	601.68	602.01	602.48	602.67	602.68	603.09	

therefore be written as

$$\widehat{X}_t = 0.830X_{t-1} - 1.324\varepsilon_{t-1} + 0.404\varepsilon_{t-2} + 0.529\varepsilon_{t-3} - 0.827\varepsilon_{t-4} + 0.634\varepsilon_{t-5}$$

Step 4: Diagnostic Checking

Goodness of fit for time series models involves testing if the model residuals form a white noise process. It is through diagnostic checks that a model can be declared statistically adequate and thereafter can be used to forecast. According to [14], if the diagnostic tests fails a new process (cycle) of identification, estimation and diagnosis is done until the best fit model is found.

The Plots of ACF, Normal Q-Q and Histogram of Residuals show that the residual are a white noise process. Thus, diagnostic check for an ARIMA (1,1,5) model in **Figure 3** indicates that the model is good (best fit).

The results in **Table 4** show that the ARIMA (1,1,5) model performed better than the SES and HWES models on FDI data for Zambia due to the minimal error. Hence, this model was picked for forecasting.

4. Forecasts

Forecasting results plays a vital role to policy makers in creating good policies and coming up with suitable strategic plans on FDI. R output of ARIMA (1,1,5) forecasts for the next 10 years of annual net Zambia's FDI's inflow is shown in **Table 5**.

Table 5 shows ten year forecasts for FDI using ARIMA (1, 1, 5). Trajectory of

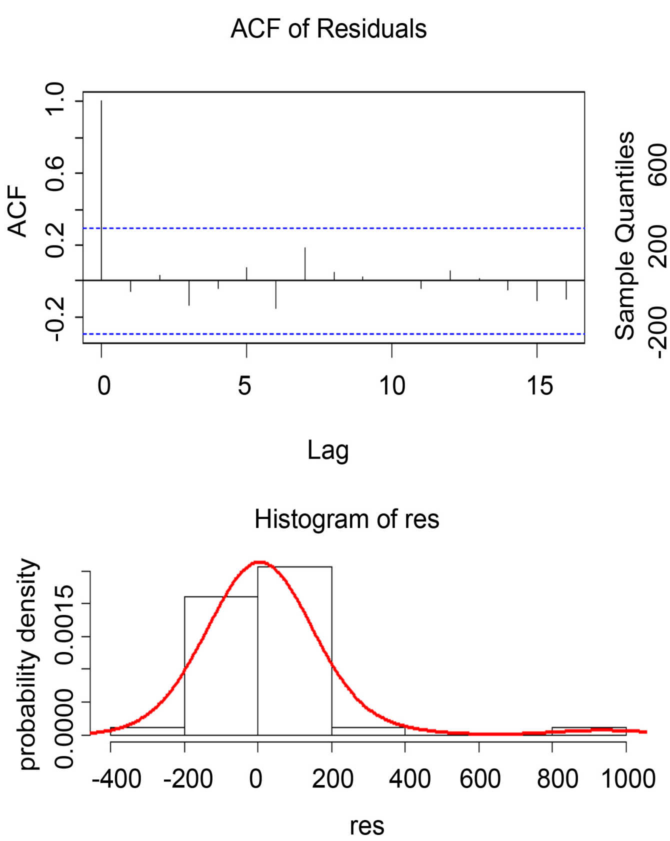

Figure 3. Plots of ACF, Normal Q-Q and histogram of residuals.

forecasts in the period from 2014 to 2024 is shown in Figure 4. Forecasting results give a gradual increase in annual net FDI inflows of about 44.36% by 2024.

5. Discussion

Forecasting is key to every field of science. ARIMA (1, 1, 5) can be used to forecast the annual net inflows of FDI to Zambia. This model can be used for both

Table 3. Estimate of ARIMA (1,1,5).

Variable	Coefficient	*p*-value
Constant		
AR (1)	0.830	0.000000036
MA (1)	−1.324	0.000000000060
MA (2)	0.404	0.10
MA (3)	0.529	0.019
MA (4)	−0.827	0.0013
MA (5)	0.634	0.00015

Note: $^*p$ value < 0.05.

Table 4. Measure of the errors for the three models.

Measure of accuracy	SES	HWES	ARIMA (1, 1, 5)
MPE	−4.573	266.981	165.380
RMSE	213.696	133.025	76.74
MAE	111.912	2.659	−9.478
MAPE	45.413	47.617	39.960
MASE	0.929	1.103	0.637

Table 5. R output of ARIMA (1, 1, 5) forecasts for the next 10 years.

Year	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2015	2353.979	2135.315	2572.644	2019.561	2688.398
2016	2442.925	2194.337	2691.512	2062.743	2823.106
2017	2757.575	2482.147	3033.004	2336.343	3178.807
2018	2665.618	2304.685	3026.552	2113.619	3217.618
2019	2806.015	2415.494	3196.537	2208.764	3403.267
2020	2922.545	2477.628	3367.463	2242.103	3602.988
2021	3019.266	2500.801	3537.732	2226.341	3812.191
2022	3099.545	2495.403	3703.686	2175.590	4023.499
2023	3166.176	2469.682	3862.670	2100.980	4231.372
2024	3221.481	2429.585	4013.377	2010.380	4432.581

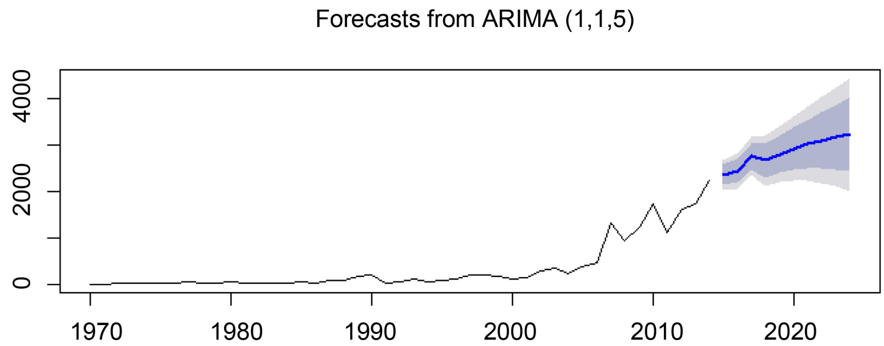

Figure 4. *R* output of ARIMA (1,1,5) forecasts for the next 10 years.

short and long term forecasting. Best strategies can only be created with accurate forecasting results. Studies have shown that FDI affects the growth of GDP. Therefore, the importance of FDI is acknowledged world over. FDI also helps diversify the country's economy (through job creation and increase in productivity), increase export in host country, improve efficiency and have technological spillovers on the already existing firms.

6. Conclusion

Three models of univariate time-series analysis were considered in this study: SES, HWES and ARIMA models. The best fit of the three models used in this study was picked based on the model indicating minimum errors. The ARIMA (1,1,5) showed smallest error than that of the SES or HWES models. Forecasting results give a gradual increase in annual net FDI inflows of about 44.36% by 2024. Policy makers use accurate forecasts to come up good policies. Therefore, the Zambian government should use such forecasts in formulating policies and making strategies that will promote FDI industry. Future research should go further and consider non-linear models such as Autoregressive Conditional Heteroscedasticity (ARCH), Generalized Autoregressive Conditional Heteroscedasticity (GARCH).

Acknowledgements

The authors are thankful to Zambia Development Agency (ZDA) for providing time series data on FDI. Many thanks also go to the Dean, School of Science, Engineering and Technology Dr Douglas Kunda for the encouragements. Not forgetting Mulungushi University for making it possible through provision of resources to come up with this research work. Also many other colleagues who made good comments on this paper.

References

[1] Fried, R. and George, A.C. (2014) Exponential and Holt-Winters Smoothing. International Encyclopedia of Statistical Science, Springer, Berlin Heidelberg. http://ceur-ws.org/Vol-1353/paper_13.pdf

[2] Chatfield, C. (1978) The Holt-Winters Forecasting Procedure. *Applied Statistics*, **27**,

264-279. https://doi.org/10.2307/2347162

[3] Box, G.E. and Jenkins, G.M. (1994) Time Series Analysis: Forecasting and Control. Prentice Hall, Englewood Cliffs. http://bookmarksland.com/download/time-series-analysis-forecasting-and-control-4th-edition-free.pdf

[4] Fu, J. (2000) Institutions and Investments: Foreign Direct Investment in China during an Era of Reforms. University of Michigan Press, Ann Arbor. https://www.press.umich.edu/16284/institutions_and_investments https://doi.org/10.3998/mpub.16284

[5] Hayter, R. (1997, Ongoing) The Dynamics of Industrial Location: The Firm, the Factory and the Production System. Notes of the Book

[6] Abdul Hasan, S. and Al-Samarrai, H. (1998) A Seminar on Foreign Investment Justification and Notification. Bayt al-Hikma, Baghdad. http://www.ijmbs.com/Vol5/3/8-abhijit-biswas.pdf

[7] Borensztein, E., De Gregorio, J. and Lee, J.-W. (1998) How Does Foreign Direct Investment Affect Economic Growth? *Journal of International Economics*, **45**, 115-135. https://doi.org/10.1016/S0022-1996(97)00033-0

[8] Tybout, J.R. (2000) Manufacturing Firms in Developing Countries: How Well Do They Do and Why? *Journal of Economic Literature*, **XXXVIII**, 11-44. https://doi.org/10.1257/jel.38.1.11

[9] Markusen & Venables (1999) Foreign Direct Investment as a Catalyst for Industrial Development. *European Economic Review*, **43**, 335-356. https://doi.org/10.1016/S0014-2921(98)00048-8

[10] ZDA (2016) Zambia Development Agency, Lusaka, Zambia.

[11] BOZ (2013) Bank of Zambia Annual Report, Lusaka, Zambia.

[12] C.S.O. (2015) National Accounts. Statistical Bulletins, No. 2, 3, 5 and 7, Lusaka, Zambia.

[13] Tularam, G.A. and Saeed, T. (2016) Oil-Price Forecasting Based on Various Univariate Time-Series Models. *American Journal of Operations Research*, **6**, 226-235. https://doi.org/10.4236/ajor.2016.63023

[14] Siluyele, I. and Jere, S. (2016) Using Box-Jenkins Models to Forecast Mobile Cellular Subscription. *Open Journal of Statistics*, **6**, 303-309. https://doi.org/10.4236/ojs.2016.62026

[15] Dobre, I. and Alexandru, A. (2008) Modelling Unemployment Rate Using Box-Jenkins Procedure. *Journal of Applied Quantitative Methods*, **3**, 156-166.

[16] Wei, W. (1990) A Time Series Analysis: Univariate and Multivariate Methods. Addison Wesley Publishing Company, Inc., New York.