BINDURA UNIVERSITY OF SCIENCE EDUCATION

FACULTY OF SCIENCE

DEPARTMENT OF MATHEMATICS

FORECASTING FOREIGN DIRECT INVESTMENT TO ZIMBABWE: A TIME SERIES ANALYSIS

BY

SIBANDA METHEMBENI M

B1852015

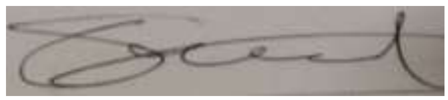
A RESEARCH PROJECT SUBMITTED TO THE DEPARTMENT OF MATHEMATICS IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF HONOURS BACHELOUR OF SCIENCE DEGREE IN STATISTICS AND FINANCIAL MATHEMATICS

SUPERVISOR: MR. K. BASIRA

2022

APPROVAL FORM

The undersigned certified that they have supervised and recommended to Bindura University of Science Education for acceptance of dissertation entitled '**Forecasting Foreign Direct Investment To Zimbabwe: A Time Series Analysis**' submitted in partial fulfillment of a Bachelor of Science Honors Degree in Statistics and Financial Mathematics.

SIBANDA METHEMBENI M		**19 JUNE 2022**
B1852015	**Signature**	**Date**

Certified by

Mr. K. BASIRA	**……………..**	**……………..**
Supervisor	**Signature**	**Date**

Certified by

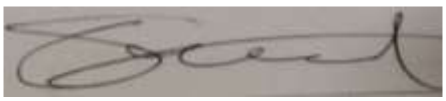
Dr T. MAPUWEI	**………………**	**……………….**
Chairperson	**Signature**	**Date**

DECLARATION

I Sibanda Methembeni M hereby declare that this submission is my own work apart from the references of other people's work which has duly been acknowledged. I hereby declare that, this work has neither been presented in whole nor in part for any degree at this university or elsewhere.

Author: Sibanda Methembeni M

Registration Number: B1852015

Signature: 

Date: 19 June 2022

DEDICATION

This research project is dedicated to my late parents Mr J Sibanda and Mrs S Sibanda.

ACKNOWLEDGEMENTS

My first acknowledgement goes to my supervisor, Mr. K. Basira for his academic support, who by his commitment and patience contributed to the successful compilation of this project. I also wish to express my gratitude to Ms. P. Hlupo and all staff members of Mathematics and Physics department for their constant encouragement and the moral support given throughout the course. My family were also of great importance; these include Dr N.N Sibanda and his wife, my late brother Mr. B Sibanda and his wife, Bekithemba Sibanda, and of course my friends for their encouragement. I dedicate this report to them for the generous patience, technical and financial support all through my life. May the highest Lord bless them.

ABSTRACT

This thesis presents an empirical study which attempts to model and forecast time series data of Zimbabwe net Foreign Direct Investment spanning from 1970 to 2021, yearly time series data was used. This study is built upon the fundamental that the researcher is tempted to feel perched among the pioneers of research of this kind, narrowing the gap on FDI, by assessing the performance of the ARIMA and GAM models in terms of accurate predictions of inflows. Probability sampling was applied for data collection and the Box- Jenkins ARIMA methodology was applied for forecasting. The diagnostic checking has shown that ARIMA (0, 1, 1) model as the optimal model to forecast FDI in Zimbabwe based on the AIC. The ADF test also indicates that the residuals of the model are stationary, thus confirming its adequacy. The results of the ARIMA (0, 1, 1) model forecast showed that the FDI inflows will remain constant. Due to the dynamic world GAM model was utilised and the results showed that the forecasts exhibited excellent properties of being best linear unbiased estimates (blue) with least Mean Squared Error (MSE) compared to the traditional ARIMA (0, 1, 1) and they are highly recommended in this research. These forecasts will help policy makers in Zimbabwe to sustain their effort to expand the tax base, reduce red tape, and strengthen the regulatory framework to investment and also investor's friendly policies in order to attract the much needed FDI. The researcher because of the limited time zoomed on assessing two models and recommending areas for further studies that other significant models can be used which were not considered in the research to come up with the best model.

Keywords: AR, ARIMA, ADF, Foreign Direct Investment, Forecasting, GAM, MA, Zimbabwe

Table of Contents

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS AND ACRONYMS

ACF	Autocorrelation Function
ADF	Augmented Dickey-Fuller
AIC	Akaike Information Criteria
AR	Autoregressive
ARIMA	Autoregressive Integrated Moving Average
ARMA	Autoregressive Moving Average
BIC	Bayesian Information Criteria
FDI	Foreign Direct Investment
GAM	General Additive Model
GDP	Gross Domestic Product
IID	Independently Identically Distributed
KPSS	Kwiatkowski-Phillips-Schmidt-Shin
MA	Moving Average
MAPE	Mean Absolute Percentage Error
MSE	Mean Square Error
MSR	Mean Square Residual
PACF	Partial Autocorrelation Function
RMSE	Root Mean Square Error
SPSS	Statistical Package for Social Sciences
ZIA	Zimbabwe Investment Authority
ZIMSTAT	Zimbabwe National Statistics

CHAPTER 1: INTRODUCTION

1.0 Introduction

The most important cornerstones of global economic growth is Foreign Direct Investment. In Zimbabwe, investment has always been a big economic challenge. The world has been trading for a long time, and all countries have issued investment regulations in developed countries with plentiful resources. Developed countries concentrate on new markets with more job opportunities, a wider product selection, and larger profit margins (Hansen & Rand, 2006). As a result, in emerging nations, Foreign Direct Investment (FDI) has become a conflict. The goal is for FDI to complement and increase domestic investment in order to attain higher levels of economic development, as well as chances for technical innovation and access to global managerial skills and processes.

The background of the study, the statement of the problem, aim of the study and objectives of the study is explained in this chapter. It also includes the research questions, the significance of the study, assumptions of the study, as well as the limitations and the delimitations of the study. Finally, this chapter discusses how the study is carried out and concludes with a summary.

Following this excellent foundation or the introductory chapter there are four additional chapters, making this a five-chapter study. Following the second chapter, a chapter examines the historical research literature by reviewing time series analysis and its usefulness in analysing Foreign Direct Investment in Zimbabwe. The research methodology chapter, which is the third in the series, focuses on methodologies employed to achieve certain study goals. The fourth chapter on data presentation, analysis and discussion is to lay out a framework for the data process and presentation of research results. The final chapter concentrates on drawing conclusions from the research and giving recommendations based on the results.

1.1 Background of the study

The United Kingdom became a Foreign Direct Investor, "while it was a trader in 1960 and became a prominent provider of direct investment, responsible for the distribution of stock to other regions of the world" (Robert & Lipsey, 2015). Therefore, FDI is known to be the driver of economic development and is a contested issue. Both arguments in favour of FDI and arguments in opposition to FDI exist. There are arguments in favour of FDI as well as arguments against it.

Some claim that FDI has a good effect on global productivity, which helps to explain why countries have different levels of economic development, while others argue that FDI diminishes local power and, as a result, causes worldwide inequality (Velde, 2006).

"Many people suggest that FDI inflows could cover the gap between targeted investments and domestic savings", (Todaro & Smith, 2003). Furthermore, many FDIs assist the country in breaking free from a vicious cycle of underdevelopment. Todaro & Smith (2011) FDI entails more than merely assisting in the transfer of funds or creating a local industry.

The background would not have been complete if the researcher overlooked FDI trends during independence. "Immediately after gaining independence in the early 1980s, newly appointed Zimbabwean executives acquired a highly centralised, inward oriented strategy that relied heavily on FDI 70% to promote economic revitalization" (Clarke, 1980). Gwenhamo (2009) pointed out that "in Zimbabwe the total FDI inflow from the country gained independence until 1990 was due to the policies introduced and implemented that were not in line with attracting foreign investors." The lengthy procedures required for foreign investors to validate their proposals and add reasons that alarmed some investors and the ownership ban that required indigenisation in other sectors. "Government policies aimed at external investors took a new turn in 1987 following the implementation of new investment policies, strategies and codes. This has led to a significant increase in FDI following the adoption of new investment policy." (Gwenhamo, 2009).

"In 1998, foreign direct investment in Zimbabwe totaled Z$444,000,000 but by 2001, it had dropped to Z$5,400,000. Private commerce and inadequate government oversight in the economic sector were to blame for the unexpected spike in FDI inflows in 1998," (Moyo, 2013). "From 2000 to 2008, Zimbabwe has been plagued by political unrest and economic disparities. All of these factors contributed to the IMF's expulsion from the country, as well as a reduction in the country's current borrowing limit. Zimbabwe's FDI inflows were disrupted as a result of the expulsion, which left the country with substantial unemployment."(Gwenhamo, 2009).

1.2 Statement of the problem

Although widely accepted and used in statistics and econometrics, the ARIMA model is not completely reliable as it gives false values resulting in the overestimation and underestimation of Foreign Direct Investments (FDIs). Having observed such, this research seeks to open up to whichever the model is identified from AR, MA, ARMA, GAM or even long memory fractional differenced model that address the weaknesses of the ARIMA model and therefore ensures that the deviation between the actual and the forecasted values is as small as possible for the purpose of better economic planning and policy making with regards to Foreign Direct Investments.

1.3 Objectives of the study

The objectives of this study are:

1. To identify a time series model;
2. To estimate the model parameters;
3. To diagnose the model;
4. Prediction and forecasting using the model from 2022 – 2026.

1.4 Research question

In carrying out the research the following questions had to be answered:

1. Will the time series model be incorporated into Foreign Direct Investment inflow?
2. What will be the future closing inflows for FDI in Zimbabwe?
3. Will the FDI fitted model be valid or not?
4. Will the inflows of FDI in Zimbabwe be constant or not constant?

1.5 Significance of the study

This study contributes to the current literature and tries to raise awareness among potential investors in Zimbabwe. It is also essential for the reader because it can improve students' knowledge of FDI and serve as a starting point for individuals interested in learning more about FDI or similar topics. Furthermore, the research assists other students in evaluating the suitability of college-taught theories and putting them into practice.

1.6 Assumption of the study

All investments, according to the study, were quantified in US dollars. Foreign Direct Investment (FDI) inflows are considered stochastic in nature.

1.7 Delimitations of the study

Because there was only one FDI dependent variable, one time independent variable and the researcher was also cut off from the time series. The research was carried out in Zimbabwe from 1970 -2021. Because the government of this country has established a highly regulated FDI economy (70%) to boost economic growth. The study has decided not to employ additional investment.

1.8 Limitations of the study

The limitations are that there is a difference in currencies applied to this study timeframe which results in biasness of the outcome. As inflation affects time series, this research study was carried out using second hand information from the World Bank. Hence, it is ideal to note the impact of inflation on forecasting. The researcher is going to convert the currency of the Zimbabwean dollar to the American dollar (USD) using Foreign Exchange rates to meet such a limit.

1.9 Definition of terms

Foreign Direct Investment

An investment made directly into industries that are production sectors by individuals or companies from foreign countries, either by takeover or merging of firms in a targeted host country or through expansion of existing operations by opening of other branches and industries. (Gwenhamo, 2009).

Time series,

A series of numerical values obtained from consecutive periods, usually with equal intervals between them (Ruey, 2010).

Economic Growth

Refers to boost and expansion of the national Gross Domestic product, which is GDP adjusted for inflation that occurs over a period of time. It is a long term growth of economic productivity (Romer, 1986).

1.10 Summary

This chapter outlined the need of forecasting Foreign Direct Investment to Zimbabwe utilising a time series analysis indicated from the title. As explained in the statement of the problem, the key aspects which triggered this research is to assess the performance evaluation of GAM and ARIMA models in estimating the future inflows, of late FDI values used to be overestimated or underestimated due to the weakness of models applied. This made the researcher come up with the objectives that focus on identifying a time series model, whichever model identified is used to make prediction and forecasting for the next 5 years as folded by the objectives. As a result, this chapter paved a way for the following chapter which is aimed at reviewing both theoretical and empirical literature.

CHAPTER 2: LITERATURE REVIEW

2.0 Introduction

This chapter examines the empirical and theoretical literature of historic research based on time series analysis, as well its success in assessing Foreign Direct Investment in Zimbabwe from 1970 to 2021. The agenda and objectives of this research are incorporated into a model study framework that is offered. The study assessed the performance of the Generalised Additive Model with the Autoregressive model in modelling FDI inflows in Zimbabwe. According Adebiyi, Adewumi, & Ayo (2014) all agree that ARIMA models are effective at generating accurate financial projections and typically outperform models that built complexity in predictive capacities. However, based on validation tests, Peter (2017) believed that GAM greatly outperformed typical models in all risk categories in modelling. This looks to be "very challenging", prompting the researcher to clarify which status should be included in certain FDI inflow figures. The literature is to provide additional impetus for this research.

2.1 Theoretical literature review

2.1.1 Time series analysis

It is the understanding of formulas that divide a series of sequential atoms, allowing natural trends to be detected, and then finding the necessary estimation and projections. The acquisition of future values from past values involves a basic comprehension of the basic context of relevant data points in time series research AR, ARIMA, MA, GARCH, TARCH, EGARCH, FIGARCH, CGARCH and ARIMA are examples of time series models, although this research concentrate on ARIMA and GAM.

2.2 The Moving Average (MA)

It is devoid of seasonal components that have a certain amount of randomness (Pannerselvam, 2005). Moving average models offer the advantage of anticipating stocks or items with a steady demand, when there is a little trend or season, and are also useful in identifying areas of support and resistance. On the other hand, MA is unresponsive to variables that occur for a cause, such as seasonal influences and cycles. The following is the formula for calculating the Moving Average:

$$y_t = \text{c} + \mathcal{E}_t + \theta_1 \mathcal{E}_{t-1} + \theta_2 \mathcal{E}_{t-2} + \cdots + \theta_q \mathcal{E}_{t-q} \ldots \ldots (1)$$

$\mathcal{E}_t$ is white noise and this is MA(q) model (Ruey, 2010).

2.3 Autoregressive Model (AR)

George, Gwilym, & Gregory (2008) describe the autoregressive model as a stochastic process that expresses a specific series of events over time.

$$y_t = \text{c} + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \mathcal{E}_t \ldots \ldots (2)$$

This model has features of multivariate regression, having lagged values of y_t as regressors and where $\mathcal{E}_t$ is the error. This is an AR (p) model which states that past p-values determine the conditional expectation for the provision of the previous data. The benefits of the Autoregressive Model are those that can be utilised to identify how much of the time series may be described by prior results in the time series. Autoregressive models compute the polynomial likelihood of the following symbol. While this is appealing, it means they won't be able to create a distribution model with difficulty to calculate the next symbol probability.

2.4 Autoregressive Integrated Moving Average (ARIMA)

The Time series should be differentiated until it is stationary, eliminating patterns and seasonal effects. The "explanatory variables" comprised of lagged values of y_t and errors, this is an ARIMA (p, d, q) model

$$y'_t = \text{c} + \phi_1 y'_{t-1} + \cdots + \phi_p y'_{t-p} + \phi_p \mathcal{E}_{t-p} + \mathcal{E}_{t,} \ldots \ldots (3)$$

ARIMA model is considered to have a unit-root nonstationary (Ruey, 2010). ARIMA forecasting's key benefit is that it only takes data from the time series in question. First, this feature is advantageous if one is forecasting a large number of time series. Second, this avoids a problem that occurs sometimes with multivariate models. However, on the other hand, ARIMA can be limited to forecasting extreme values. While the model is adept at modelling seasons and trends, outliers are difficult to forecast for ARIMA for the very reason that they lie outside of the general trend as captured by the model.

2.5 Seasonal Autoregressive Integrated Moving Average (SARIMA)

The SARIMA model works with both seasonal and non-seasonal data. The seasonal fluctuation of adequate order is employed in this model to make the series stationary. The definition of the model is as follows:

$$(1-\phi_1 B)(1-\Phi_1 B^4)(1-B)(1-\ B^4)\ y_t = (1+\ \theta_1 B)(1+\Theta_1 B^4)\mathcal{E}_{t,} \dots \dots (4)$$

B is the backshift operator with '4' the seasonal lag (Ramasubramanian, 2015).

2.6 Generalised Additive Model (GAM)

Aside from using the correlation between values from the same time points, the general trends are modelled. The time series may be viewed as the accumulation of individual patterns. In the generalised additive model, financial market data can be modelled by introducing seasonal trends to a growing growth trend (Annalyn & Kenneth, 2017). The GAM principle is almost the same with regression besides that of summarising the outcomes of single predictions. GAM generates incredibly complicated trends which are quantified in achieving the most common smooth curves. With facts that GAMs are centred on functional forms rather than flexibility, they are not confined to linear regression that demands prediction and flexibility of outcomes to be in a linear form. Annalyn & Kenneth (2017) GAM employs a process known as back fitting to obtain the best trend line that matches the data. Back fitting is a process that repeats positive adjustments in GAM tasks of providing with the trend line that minimises guessing error.

2.7 Forecasting using the Facebook Prophet (GAM)

Is a method of predicting trends and made up of an additive model that incorporates nonlinear patterns, as well as annual, weekly seasons and it works best with a year's worth of daily periodicity data. The missing data, trend shifts and big outliers have little effect on Prophet (Taylor & Letham, 2018). The functions that make up the GAM trend findings using the Prophet package are classified as follows:

Overall Growth - To model this, a linear or logistic the pattern could be used.

Seasonal Variations - Fourier series, which is a method of measuring periodic tasks, exemplifies this. Back fitting is a method of determining the exact functionalities. One can specify whether it is expecting weekly or annual trends to exist.

Special Events - Aside from modelling recurring trends, one-time events must also be taken into account. This covers any event, such as elections, bond announcements, or the downgrade of Zimbabwe's credit rating, that might cause ripples in natural disasters. If irregular events are not taken into consideration, the GAM may misinterpret them as persistent occurrences, causing their consequences to be propagated incorrectly. To account for the ripple effect, one can optionally define windows before and after each occurrence (Annalyn & Kenneth, 2017).

2.8 Empirical literature review

Nyoni (2018) utilised the Box-Jenkins ARIMA method to forecast net FDI inflows in Zimbabwe from 1980 to 2017. The researcher gathered yearly time series data on net FDI inflows in Zimbabwe. The ADF test revealed that the data on FDI was not stationary. As a result, the research chose the lowest AIC value and then presented ARIMA (1, 1, 1) as the best model for forecasting FDI. After differencing the series, the ADF test revealed that the residuals of the ARIMA (1, 1, 1) model are stationary, validating its adequacy. The forecasted values for FDI inflows over the next two decades revealed a mediocre and lacklustre, innovative economy. Among the primary policy recommendations, the research also suggested that policymakers in Zimbabwe develop policies that attract foreign investors.

According to Henry, Elijah, Gwani, & Simon (2019), in their research time series ARIMA model for predicting Nigeria net Foreign Direct Investment, they employed the Box-Jenkins ARIMA approach. The annual data was gathered from 1972 to 2018, and the study showed that at the first difference, the series became stationary. Model diagnostic testing revealed that ARIMA (1, 1, 2) was the best model based on the AIC, BIC, and Hanna Quinn Criterion (HQ). This study performed a 20-year forecast from 2019 to 2039, and the results revealed that net FDI in Nigeria continued to expand over the forecasted period. As a result, the prediction was intended to assist Nigerian policymakers in maintaining their efforts to broaden the tax base and decrease red tape, and cement, the legal environment for investment as well as the investor’s friendly policies in order to attract the much needed FDI.

Idowu (2021) conducted research on econometric modelling and Forecasting Foreign Direct Investment inflows in Nigeria over the next decade. The study used the Box-Jenkins ARIMA model technique, and the research span was from 1970 to 2020. The correlogram shows that the Nigerian FDI net inflows were integrated into the first order. In relation to the number of significant coefficients, the greatest adjusted R-squared, the lowest volatility, and the lowest AIC, the study recommended the ARIMA (1, 1, 3). Cross validation and diagnostic testing additionally revealed that the estimated model is not only consistent, but also effective in forecasting and explaining Foreign Direct Investment inflows in Nigeria. The study's findings revealed that net FDI inflows in Nigeria are most likely to demonstrate a very modest increasing trend. The research made suggestions to policymakers and the central government of Nigeria on how to effectively accelerate and sustain greater levels of net inflows.

Nyoni & Muchingami (2018) conducted research on the Foreign Direct Investment dynamics in India using yearly time series data on net FDI inflows from 1960 to 2017. The study employed Box-Jenkins ARIMA methodology to examine the inflows. Several tests were performed, and the ADF test revealed that the data was not stationary, necessitating the need to differentiate a series. The KPSS test revealed that the results are consistent with the ADF test results. Based on the AIC, the study presented the ARIMA (1, 1, 0) model as the optimal. The diagnostic tests were explored further, and it was shown that the parsimonious model is not only stable, but also suited for studying FDI dynamics in India. The findings of the study showed that net FDI inflows are most likely to weaken over the next ten years. As a result, the study makes meaningful policy suggestions to policy makers on how to encourage and boost much anticipated FDI inflows in India.

Annalyn & Kenneth (2017) carried out a research on evaluating patterns in Daylight Saving Time (DST) website visits of individuals living in four season zones to demonstrate that Generalised Additive Models have a wide range of application in our daily lives. The Python package was used to disclose DST page visit statistics from 2008 to 2015 on the Wikipedia website. Overall growth, unusual events, and seasonal variation plots were discovered after fitting a suitable GAM. After that, it became clear that the overall page view of the DST Wikipedia article had decreased over time. People are less likely to learn of DST during weekdays according to weekly statistics. General patterns revealed that the website visit increase between the month end of March and October.

Peter (2017) did research and demonstrated how to use GAM in Microsoft R to analyse electricity use. The following data, manipulation and deception, time series plots were created, yielding two main seasonalities: daily and weekly. The maximum number was seen in these plots at 1500 hours on weekdays and decreased on weekends. The model's residual plots revealed that the normality assumptions have not been violated in any way. Forecasts for future power consumption were created utilising the fitted GAM. Based on the results, he concluded that "GAM is smoother than Multivariate Regression" (Peter, 2017).

2.9 Research issues and research gap

The foregoing literature review is useful in determining research problems and gaps that are primarily the foundations upon which the objectives of the current study are centred on. Much research has been conducted on a time series analysis of Foreign Direct Investment, although the majority of these studies have concentrated on the regional and continental levels. There has been less attention placed on the Southern area, notably Zimbabwe. Most of these studies did not examine the effectiveness of ARIMA and GAM models in forecasting FDI values; instead, researchers relied only on the ARIMA model and concluded that it is the best model estimator for FDI inflows. However, the researcher is tempted to feel perched among the pioneers of research of this kind, narrowing the gap on FDI, by assessing the performance of the ARIMA and GAM models in terms of accurate predictions of inflows in Zimbabwe. This study differs from previous research in several aspects and adds to the current literature body of knowledge in the following ways. Firstly, the current study examines patterns of FDI. Secondly, the current study aims to highlight developing countries' changing attitudes toward FDI as well as developed countries' changing attitudes toward developing countries in order to better understand their contributions to contemporary international relations and development.

2.10 Summary

This chapter examines literature on FDI and time series analysis, as well as dynamic evaluations of prior works. Following the review of literature, there are some conclusions that early research only concentrated on forecasting FDI time series analysis utilising ARIMA models. Therefore, this caused policy makers to embark on wrong decisions since these models overestimated or underestimated the results. As a result, there was a gap that aside from using ARIMA models, other better models such as MA, ARMA, GAM or even long memory fractional difference models can be used. Following this gap, the researcher is tempted to feel perched among the pioneers of research of this kind, narrowing the gap on FDI, by assessing the performance of the ARIMA and

GAM models in terms of accurate predictions of inflows in Zimbabwe. The next chapter examines the research process, including how data was acquired and how it would be examined.

CHAPTER 3: RESEARCH METHODOLOGY

3.0 Introduction

This chapter concentrates on methods implemented to meet specified study objectives. The methodology is not limited to research methods, but can also be used as a basis for strategies used in the discourse of this study and to explain why the researcher used some techniques and not another, thus making the researcher the only person who can evaluate the results of this study (Kothari, 2017). The research methodology gives a framework and tools used for data collection and the plan for data analysis (Dawson, 2013). The time series model is to be injected to unveil and predict future FDI inflows.

3.1 Research design

Refers to a key tool used by researchers as it provides guidelines on how the study should be conducted. This study pays attention to the development of quantitative research design. Thus, it allows for the study and interpretations of problems affecting economy. According to Bayai & Nyangara (2013) both have shown that it fits well with the quantitative analysis and comparisons amongst groups. The goal of the research design is to gain more control over the study and increase its validity by studying the research problem. This study chose a quantitative-descriptive research design.

3.2 Research instruments

Research tools or instruments are developed to collect data (Nieswiadomy, 2010). The researcher used a laptop to download FDI data. Tables used for data storage, Microsoft Excel package was used for viewing data and Microsoft R programming language containing statistical libraries to analyse the collected data.

3.3 Population, Sample and Sampling Methods

FDI annual data was collected from 1970-2021 and used in this research study. This study's sample size was 52 observations. The researcher in order to select a sample, applied probability sampling there by giving each sample equal chances of being randomly selected.

3.4 Data Source and Problems

The research exploited secondary data for the FDI time series from 1970 to 2020 from numerous sources because it was not difficult to find and also not complex to investigate. All variables in this study are on an annual basis. Time series data was selected because it captures variability in time rather than heterogeneity between variables. The time series is able to provide insight into what was happening between the variables of concern over the period under study.

Limitations are the restrictions of the scope of the study and may result in elbow grease in the study completion (Demaria & Kothari, 2017). Different nations might have the same GDP, which is why income distribution in each nation is not the same. However, making comparisons between these countries may seem daunting, working hours generated by a certain income level may vary, and one sided GDP numbers may require broader economic welfare measures. Global prices will also vary, which is important because the will to buy is based on the value of the property. To address this issue, GDP figures can be reproduced in terms of consumer buying behaviour.

In general, collecting and compiling FDI statistics is difficult. Therefore, extra vigilance and concentration are especially needed when working with FDI statistics to analyse trends in all categories and to measure as well as deducing the effect of MNCs to the host economy, as some statistical data could easily mislead policymakers. In order for Companies in the domestic economy to be recognized as FDI, they must transfer funds to MNCs through a proxy country. Funds are sometimes transferred to a foreign assortment, and are transferred directly from the source. This allows companies to benefit from tax. The process is called the round-trip, and can count a large part of FDI. A major problem with the use of FDI data is that no inflows developed with this procedure, which unfairly increases figures. This impact is reflected in China's state of affairs since the new millennium. Although operating capital accounts, this approach may be as a result of business taxation on MNCs.

Data gathered for various purposes with another person on behalf of a user is secondary (Nieswiadomy, 2010). Second hand data may lead to biased results, as of this research it is confirmed. The data was downloaded from World Bank. Data was stored on a laptop in Excel format.

3.5 Data presentation and Analysis procedures

3.5.1 Components of the Time series

The most crucial step in choosing the best model and prediction procedure is to take into account the type of patterns that the data exhibits or shows from the plots. As a result, in order to use the most accurate methods, these steps include:

Trend - is a "long-term progressive movement" with no calendar agnate and asymmetrical impact, and is a clear level of integration, " (Australian Bureuau of statistics, 2008). Consumers are the result of time line data, showing up-and-down trends over a period of years (Blog, 2008). “Trends may show a downward or upward spiral pattern (Adhikari, 2014).

Cyclic - Makridakis, Hyndman, & Wheelwright (2007) agreed that the cyclic pattern occurs when the data display an increase and decrease in a non-fixed cycle, and the cycle does not repeat periodically (Pannerselvam, 2005).

The Seasonal - Blog (2008) stated that "Seasonal variation is a temporary variation in a series of times that occurs at certain times of the year.”

Irregular - the unfair part (sometimes also known as the residue) is the only left after the season parts and series style are limited and remote (Australian Bureuau of statistics, 2008). Unusual variation is the variability of the short term series, which is flexible and does not follow the pattern of the event (Blog, 2008).

3.6 Box Jenkins methodology

The Box–Jenkins technique utilises ARIMA models to identify the optimal model needed to fit historical values in time series analysis. Below is the representation approach of Box-Jenkins:

Model identification: The beginning stage of the Box – Jenkins model is determining if the time series is stationary and whether substantial seasonality has to be modelled. The run sequence graphic can be used to check for stationarity. The run sequence graphic should be consistent in terms of position and scale. To achieve stationarity, Box and Jenkins offer a differencing strategy. Fitting the curve and deleting the values present in the real data, on the other hand, may be employed in the context of Box – Jenkins models. After adjusting for stationarity and seasonality, the next step is to determine the order (i.e. p and q) of the autoregressive and moving average components.

Model estimation: Numerically estimating the solutions of nonlinear equations is used to estimate the parameters for Box–Jenkins models. As a result, statistical software designed to handle the approach is common almost all modern statistical packages have this capability. Nonlinear least squares and maximum likelihood estimation are two basic techniques to fitting Box–Jenkins models. In most cases, the maximum likelihood estimate is the recommended method.

Model diagnostics: The residuals should be drawn from a stationary distribution satisfying homoscedasticity condition, and should be white noise. These assumptions should be satisfied to show that the residuals produced best fit to the data. Statistical visuals of the residuals are one approach to see if the Box–Jenkins model residuals obey the assumptions. Another option is to look at the Box–Ljung statistic's value. Below is the summary of Box-Jenkins approach:

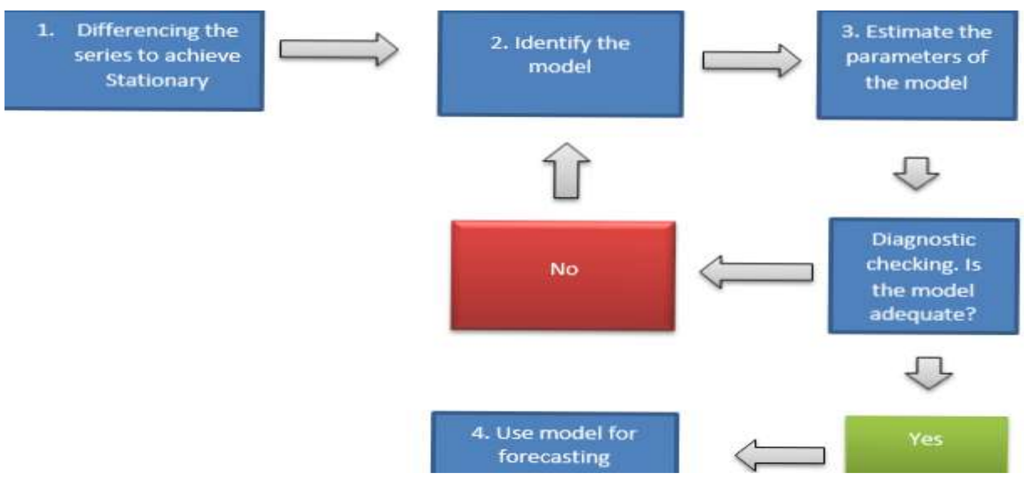


Figure 3.1 Box- Jenkins summary approach

3.7 Diagnostic checking

3.7.1 Testing for stationarity: Augmented Dickey Fuller

Raw data is tested to determine which models are selected to represent a series. Autocorrelation functions and partial autocorrelation functions are part of the series needed to achieve this. Data differencing can be applied if raw data does not stop as well as ADF is applied to validate stationarity. The greater the rejection of the hypothesis that there is a unit root at whatever degree of confidence, the more negative it is.

3.7.2 Kwiatkowski–Phillips–Schmidt–Shin (KPSS)

Compare the alternative of a unit root to the null hypothesis, which implies that the series is stationary. We infer that the time series is non-stationary if the KPSS value is lower compared to 0.05, and we fail to reject the null hypothesis if the p-value is greater than 0.05.

3.7.3 Akaike Information Criteria

The Akaike information criteria (AIC) is used to obtain the order p of the AR process (Ruey, 2010).

$$AIC = \frac{-2}{T}\ln(\text{likelihood}) + \frac{2}{T}(\text{number of parameters}) \ldots\ldots (5)$$

Maximum likelihood estimates are used to evaluate the likelihood function, and T is the sample size (Ruey, 2010). The AIC is reduced to the Gaussian AR (ℓ) model.

$$AIC(\ell) = \ln(\overline{\delta_\ell^2}) + 2\frac{\ell}{T} \ldots\ldots (6)$$

$\overline{\delta^2_\ell}$ is the maximum-likelihood estimate of δ_a^2 and T is the sample size, and is the variance of, a_t respectively (Ruey, 2010).

3.8 Forecasting performance criterion

One of the primary goals of time series models is to predict or anticipate the future, however unacknowledged it may be. One of the objectives of creating a time series model is to be able to predict future FDI inflows for that series (Cryer & Chan, 2008).

If the model is well-chosen and its parameters are well-measured, it performs well when used to generate forecasts, and users of predictions balance the benefits and downsides as time goes on (Makridakis et al., 2007). In this case, the researcher used GAM and ARIMA models to produce ZSE data predictions.

3.9 Summary

This chapter describes the study's research methodology. In trying to answer research questions, a detailed account of all actions carried out during the study was written. The chapter moved on to discuss data display and data analysis methods, which is covered in the following chapter. The next section concentrates on data representation, interpretation, and analysis.

CHAPTER 4: DATA PRESENTATION, ANALYSIS AND DISCUSSION

4.0 Introduction

To complete this study effectively, data was analysed to evaluate the hypothesis and respond to the research questions mentioned earlier in the preceding chapters, Data is presented in a descriptive manner. This chapter covers the analysis, presentation and interpretation of the results of the study. A retrospective time series analysis was performed to obtain meaningful data for the research objectives. The ultimate goal of the study is to test and compare the performance of the GAM and ARIMA models in modelling Foreign Direct Investment inflows in Zimbabwe.

4.1 Descriptive statistics

Table 4.1 Descriptive Table

Min.	1st Qu.	Median	Mean	3rd Qu	Max.
-3.0506684	3371701	28980000	104899760	125715000	744637199

Descriptive statistics show minimum, maximum, and quartile, median, and mean values, and reflect the outliers of the 52-year FDI data surveyed. The data was imported from Excel into R in CSV format and was called "data" to be displayed as time series data.

4.2 Data analysis and discussion

Model specification and identification

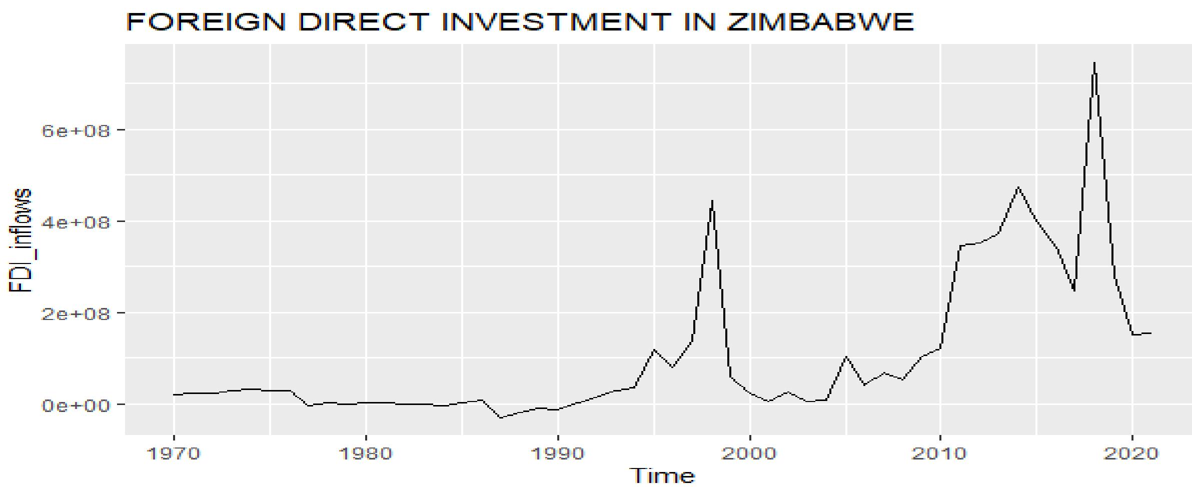


Figure 4.1 Time series Plots of Raw Data

Trade barriers and heavy restrictions had a detrimental influence on this era from 1980 to 1990. The FDI trend was nearly the same from 1980 to 1986, with a large increase the following year, 1987. This trend represents the demise of FDI from 1987 to 1991. Towards the end of the 1980s, some provisional liberalisation measures were put in place, culminating in the Economic Structural Adjustment Programme (ESAP) from 1991 to1995. The goal was to promote FDI inflows through the use of foreign exchange markets. FDI fell in 1996 before skyrocketing to $444,000,000 in 1998. The decrease from 1999 to 2000 was due to international investors who had lost faith investing in Zimbabwe.

There was a minor decrease in FDI inflows in 2008, followed by a rise in 2009 and 2010, the most likely owing to greater investment in our economy. FDI inflows increased significantly in 2011, 2012, 2013, and 2014, respectively. The inflows during this era nearly doubled the largest annual FDI inflows in Zimbabwe, which occurred in 2010. In 2015, there was a modest decrease in FDI inflows, but it was still higher than in 2013. In 2016, the country received only 33USD $.3 million in FDI inflows. From 2017 to 2020, FDI continued to shrink as a consequence of economic and political variables that exacerbated the country's economy by decreasing the rate of FDI in the country.

4.2.1 FDI Trend in Zimbabwe,

The stationarity test is performed utilising ADF, which examine the origin of the unit root. To determine stationarity, the significant value or p-values is utilised. The KPSS Test is also used in the study to determine the presence of a trend (Henry et al., 2019)

Table 4.2 KPSS Test for Trend Stationarity

KPSS Test for Trend Stationarity
data: DataFrame
KPSS Trend = 0.12089, Truncation lag parameter = 10, p-value = 0.09649

The KPSS test revealed that the significant value is greater than the printed value, implying that the series is stationary. As a result, the obtained results explain that the data is consistent with ADF, as agreed by Nyoni & Muchingami (2018) when they conducted the test in time series data to avoid spurious regression problems and violet of the assumption of the Classical Regression Model.

4.2.2 Stationarity test

Table 4.3 ADF test for Level Stationarity

Augmented Dickey-Fuller Test data: DataFrame
Dickey-Fuller = -2.698, Lag order = 5, p-value = 0.2936 alternative hypothesis: stationary

The null hypothesis was rejected because the p-value in the ADF test was greater than 0.05 and the time series variable was determined to be stationary. This means that the requirements of homoscedasticity have been met, as demonstrated by (Henry et al., 2019). The ADF test conducted by Henry, Elijah, Gwani, & Simon (2019) produced the same meaning of results as in this study, that the FDI series became stationary after the first difference. This means that the FDI series is integrated into order one, which is I (1) as agreed by (Nyoni, 2018).

4.2.2 Autocorrelation function (ACF) and Partial autocorrelation function (PACF)

If an ACF does not die off quickly, it is assumed that the data is not stationary. Fig below indicated that the ACF plot exhibits the correlation on the most lags and that the plot has not died out quickly.

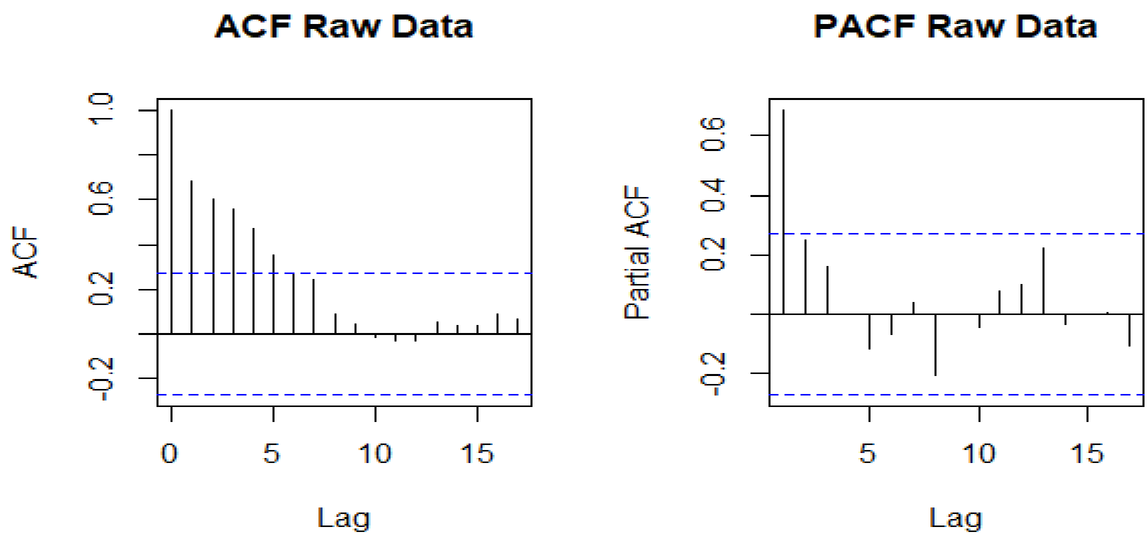


Figure 4.2 ACF and PACF Plots of Raw Data

4.2.2 Model Identification

The ARIMA (p, d, q) models, where d is the integrated part showing the number of differenced times when p is calculated from the AR model and q from MA. The ACF and PACF as part of differenced data are used to calculate p and q values, and the following models are rated using the terms "aicc", "aic", and "bic":

Table 4.4 ARIMA Modelling

ARIMA(0,1,0)	:2050.936
ARIMA(0,1,0) with drift	:2053.082
ARIMA(0,1,1)	:2042.317
ARIMA(0,1,1) with drift	:2044.313
ARIMA(0,1,2)	:2044.459
ARIMA(0,1,2) with drift	:2046.412

ARIMA(0,1,3)	:2046.676
ARIMA(0,1,3) with drift	:2048.81
ARIMA(0,1,4)	:2048.313
ARIMA(0,1,4) with drift	:Inf
ARIMA(0,1,5)	:2049.628
ARIMA(0,1,5) with drift	:Inf
ARIMA(1,1,0)	:2045.672
ARIMA(1,1,0) with drift	:Inf
ARIMA(1,1,1)	:2044.482
ARIMA(1,1,2)	:2046.787
ARIMA(1,1,3)	:2049.108
ARIMA(1,1,3) with drift	:2051.341
ARIMA(1,1,4)	:2050.533
ARIMA(2,1,0)	:2044.879
ARIMA(2,1,1)	:2046.609
ARIMA(2,1,2)	:2048.619
ARIMA(3,1,0)	:2045.98
ARIMA(3,1,1)	:2048.122
ARIMA(4,1,0)	:2047.811
ARIMA(4,1,1)	:2050.387
ARIMA(5,1,0)	:2050.386
Best model: ARIMA(0,1,1)	

The AIC criterion was used to choose the best model, and the model with the lowest AIC value was chosen. Also, as suggested by Idowu (2021), depending on the percentage of significant coefficient, the greatest coefficient of determination, the lowest volatility, and lowest AIC, the study recommends ARIMA (1, 1, 3). Therefore, as a result, the researcher used the same techniques to choose the ideal and appropriate model for this analysis, and it was noted that ARIMA (0, 1, 1) had the lowest AIC value after estimating all the parameters and possible models. Moreover, the ARIMA (0, 1, 1) meets the properties of the white noise process because the residuals have been proved to be stationary. Although numerical potential models may adequately reflect the data, the concept of parsimony is used to pick a model that adequately fits the data based on the number of parameters, (Nyoni, 2018).

4.2.3 Model Structure

The chosen form of the FDI model in Zimbabwe is in the form;

$$X_t = -0.5103\mathcal{E}_{t-1} + \mathcal{E}_t$$

The findings show that the value of FDI inflow has a negative coefficient (-0.5103), which implies that the Foreign Direct Investment inflows can contribute to the country's economic and social growth. Also, the standard error (s.e) obtained is 0.1319 with sigma^2 estimated as 1.345e+16. According to Idowu (2021), the obtained results are consistent with our expected outcomes, where FDI inflows will lead to economic growth.

4.3 Diagnostic Checking

The residuals were examined to see if the ARIMA model prediction was satisfied. The Ljung-Box test, ACF and PACF plots, histograms, and density charts were also done. The diagnostic analyses shown in Figure 4.3 utilising the residual plot demonstrate that the structure of FDI appears to be constant, although certain residuals deviate from the mean zero and constant variance. Coincidentally, Idowu (2021) discovered that diagnostic testing revealed that the estimated model is not only consistent but also appears to agree with the results of this study. Therefore, Idowu (2021) added on to say that the estimated model is also effective in forecasting and explaining Foreign Direct Investment inflows in Nigeria. As a result, the researcher is confident that ARIMA (0, 1, 1) is effective in forecasting FDI inflows in Zimbabwe.

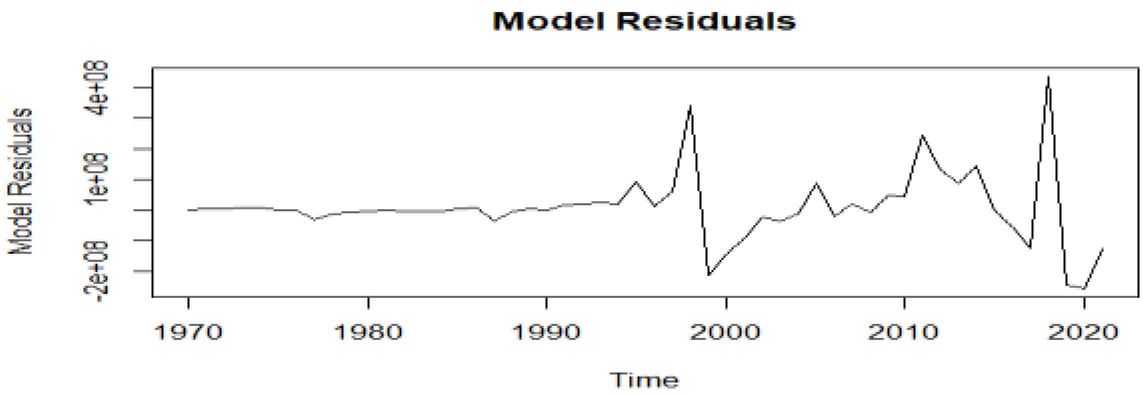


Figure 4.3 Model Residuals Results

The autocorrelation (ACF) plots show that, with the exception of lag 1, sample autocorrelation falls within the 95% confidence interval for the first 15 lags. The PACF of the residuals shows that the residuals' autocorrelation is all zeros, implying that all lags are uncorrelated. This leads to the conclusion that the residuals have a constant variance, the selected model ARIMA (0, 1, 1), and the real mean of the residuals is approximately zero. As a result, the chosen model meets all of the model assumptions. Also, as demonstrated by Henry, Elijah, Gwani, & Simon (2019), the ACF and PACF plots of the residuals from the ARIMA (1, 1, 2) reveal that all correlations are within the tolerance limit, implying that the residuals behave similarly to white noise;

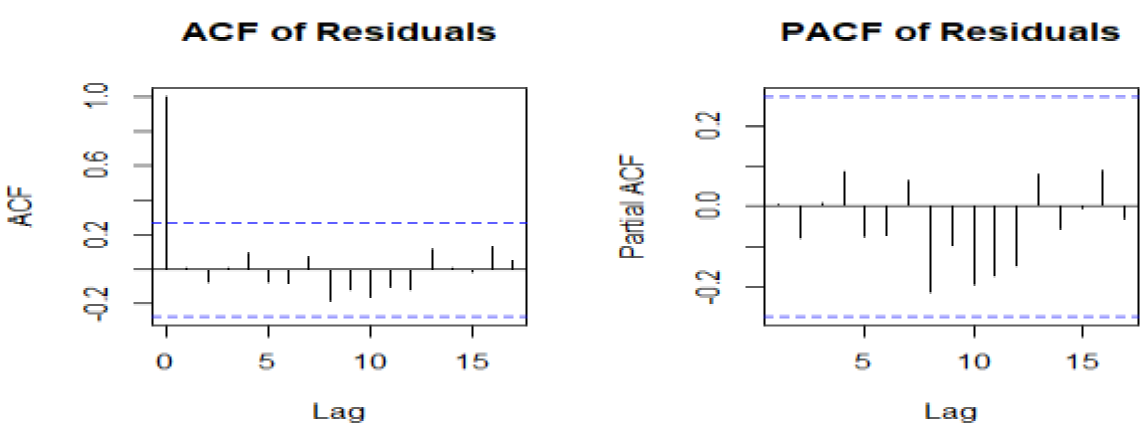


Figure 4.4 Testing for Independence of Residuals

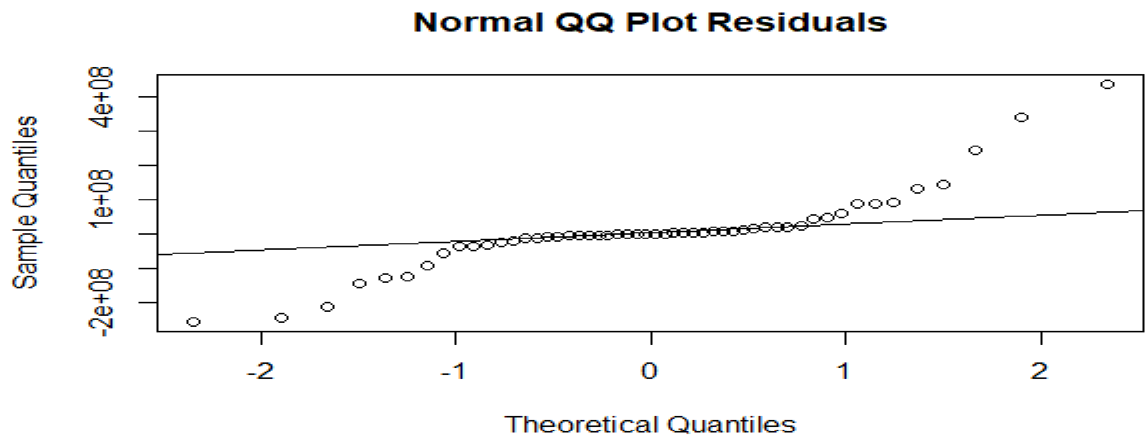


Figure 4.5 Testing for the Normality plot of Residuals

Figure 4.5 depicts a normal Q-Q plot, which illustrates that the normal distribution gives a satisfactory fit for this model since the extreme values somewhat tail off and the majority of the values are on the line, indicating that the residuals are normally distributed.

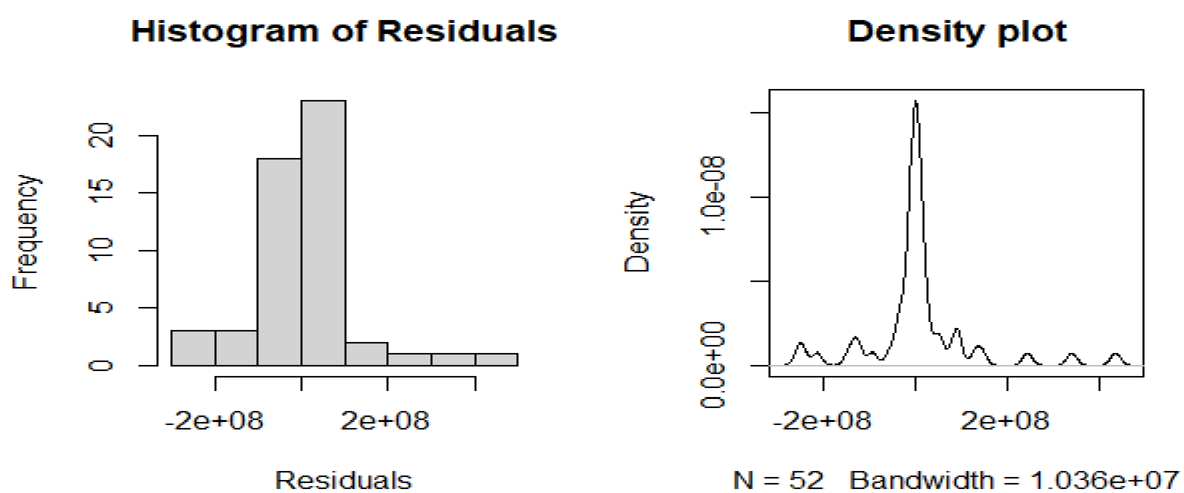


Figure 4.6 Testing for Independence of Residuals

The histogram and density chart in Figure 4.6 reveal that the residuals have a bell-shaped distribution, indicating that they follow a normal distribution. Furthermore, the mean is zero, indicating that it qualifies as white noise and the homoscedasticity condition is met (Henry et al., 2019).

Table 4.5 Box-Jenkins test

	Box-Ljung test	
X-squared = 0.0013133	df = 1	p-value = 0.9711
X-squared = 1.1652	df = 5	p-value = 0.9482
X-squared = 6.9064	df = 10	p-value = 0.7343
X-squared = 9.7423	df = 15	p-value = 0.8356
X-squared = 23.608	df = 20	p-value = 0.2599
X-squared = 25.417	df = 30	p-value = 0.7045

Table 4.5 reveals p-values for all the lags which are greater than 5%, showing that the residuals are stochastic and the model gives a satisfactory fit to the Foreign Direct Investment data. Meanwhile, because the ARIMA (0, 1, 1) model meets all of the necessary assumptions, it can be stated that the model gives a satisfactory fit of FDI data. Henry, Elijah, Gwani, & Simon (2019) highlighted that the Ljung-Box test results return a relatively high p-value, concluding that the model is appropriate. It can also be noted from this analysis that several lags were applied, and the p-values remained above 5%. Hence this implies that the ARIMA (0, 1, 1) is also adequate. The following forecasting model would be created using the parameter estimation in Table 4.4.

4.4 Model validation

Model fitting that is ARIMA (0, 1, 1) using training data has now become the most common and efficient practice for evaluation of a model's performance on test data where comparisons on the test data use dissimilar forecast horizons. The 1-step errors on the test data were obtained and used as training data to estimate parameters. The concepts of data preceding were applied to each other, both on training and test data, when computing forecasts. The researcher commenced by correlating anticipated to actual time series values, which helped to determine the effectiveness of FDI estimates. Figure 4.7 summarises the results.

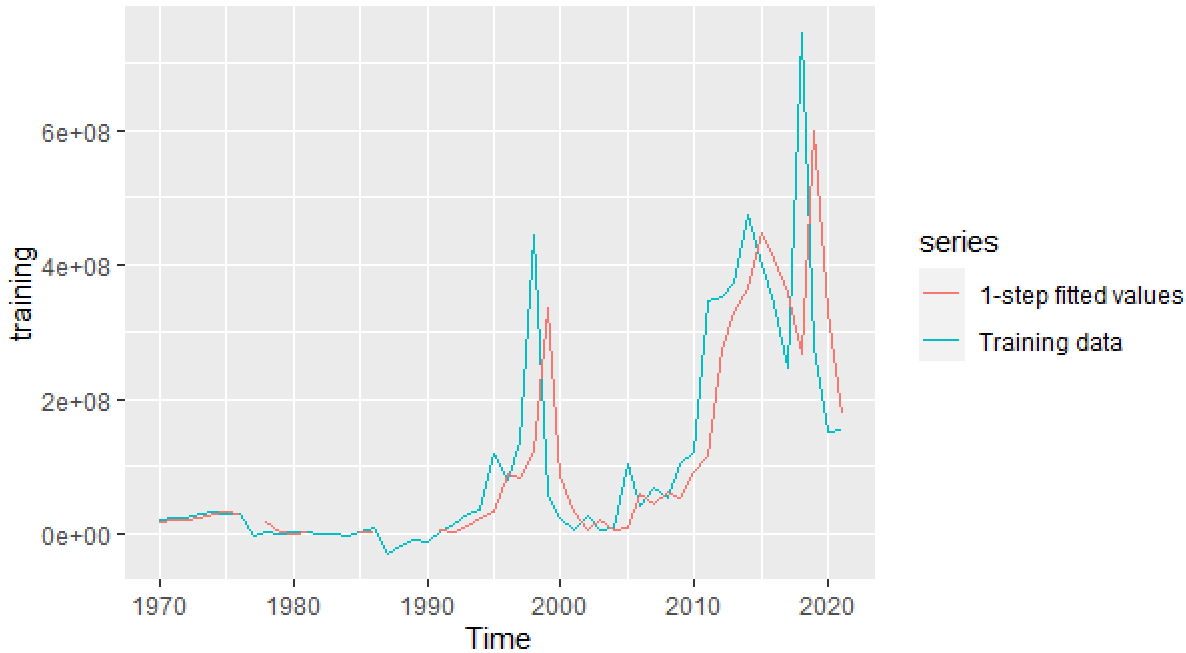


Figure 4.7 Model validation 1-step fitted values

Overall, this forecast corresponds closely with genuine numbers, indicating a general upward trend and is important to assess the validity of our forecasts. Previous researchers have demonstrated that in many cases, the best model does not prove to be the best predictor, which has contributed to additional research to test the accuracy of the model using tests such as Mean Square Error (MSE), Root Mean Square Error (RMSE), and other similar tests performed in accordance with the model (Idowu, 2021).

Table 4.6 ARIMA Model accuracy measure

Training set error measures:

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	7790285	113742389	63769893	-143.33	278.393	1.04757	0.00488

Table 4.6 shows the accuracy measurement for the ARIMA (0, 1, 1) on the FDI inflows in Zimbabwe. Applying the scale measurement unit RMSE of one step ahead forecast yields a value of 1.1E+08, which is very high as it is greater than 0. The MAE forecast is 6.4E+07 away from

the actual FDI inflows, signifying that the model has overestimated the accuracy percentage. Smaller values of statistics would generally mean the forecast is fairly excellent, but our model seems to have a very high inequality coefficient of MAPE, which is roughly 278.4, implying that our model does always have forecasting capabilities (Nyoni & Muchingami, 2018). As a result, there is a need to consider other forecasting models such as GAM.

4.5 Forecasting

The key factor or aim in international decision-making has been made easier and it is engineered through forecasting. Therefore, playing a pivotal role in providing a clear, broad knowledge of future events based on previous events and current data flow. Box and Jenkins (1976) established the foundation for economic policies relating to company planning and industrial processes, stock control, and production.

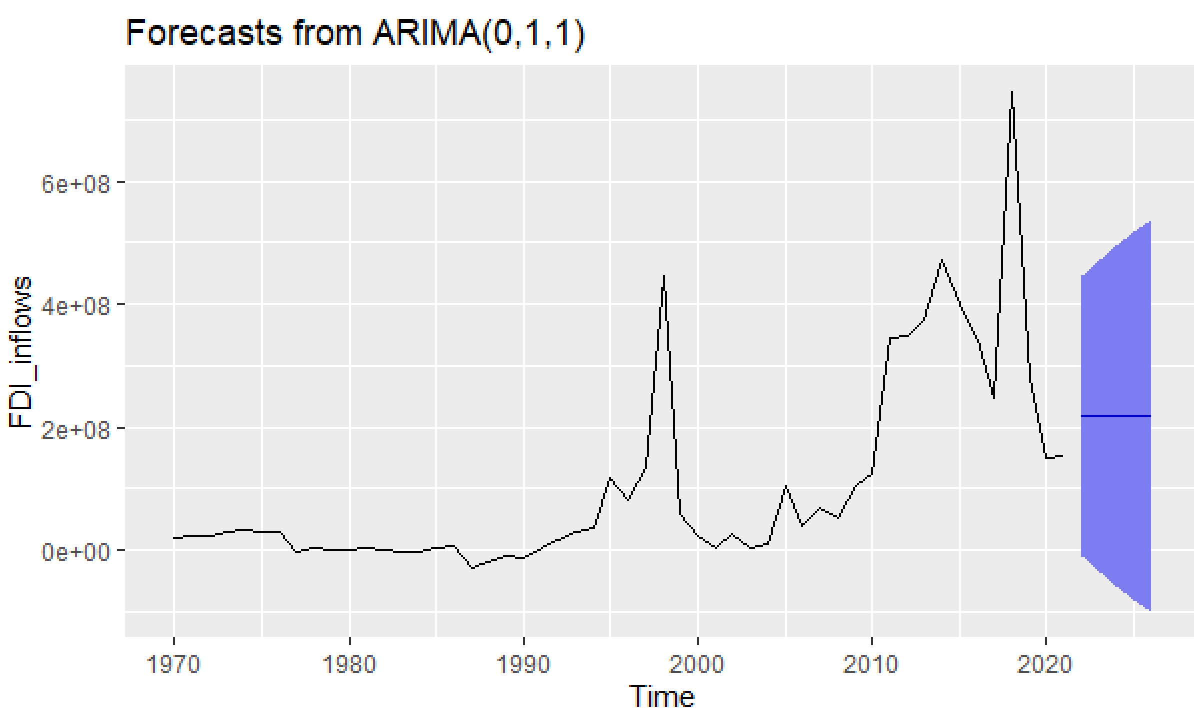


Figure 4.8 Forecasted values plot

The above diagram shows the blue line's forecasted values. The researcher observed that FDI had been increasing constantly into the future from 2021 to 2025. The forecasted value is 2.2E+08 for the next five years.

Table 4.7 Forecasted FDI up to 2026

Year	Point Forecast	Lo 95	Hi 95
2022	217744637	-9601252	445090526
2023	217744637	-35395219	470884493
2024	217744637	-58793643	494282916
2025	217744637	-80361139	515850412
2026	217744637	-100470207	535959481

It's indeed normal for us to lose confidence in our values as we predict more into the future. This is demonstrated by the ARIMA (0, 1, 1) model confidence intervals, which increase broader as we approach further into the future, as seen by the blue shade surrounding the blue line (Ruey, 2010).

4.6 GAM Model

In this section, the researcher carried out a step-by-step procedure for locating the optimum model. The goal is to assess the performance evaluation of ARIMA and GAM in forecasting Foreign Direct Investment data. The data was uploaded to R-Studio and named “tsc.” Also, Peter (2017) demonstrated his study using GAM in Microsoft R to analyses electricity use. Since Prophet does not read Date as the independent variable and FDI inflows as the response, the researcher stored Date in “ds” and FDI inflows in “y”. Therefore, this new data frame was kept as “df” this is shown below:

Table 4.8 FDI Inflows data

Ds	y
1 1970-12-31	18670000
2 1971-12-31	21840000
3 1972-12-31	23490000
4 1973-12-31	30490000
5 1974-12-31	33350000
6 1975-12-31	28870000

4.7 Model Identification and Fitting

To determine and fit the model specified by a new Prophet Object *model1*, any settings for prediction techniques are therefore passed into the constructor. We subsequently invoke its fit method, passing in the historical data frame. Prediction and forecasting are made in the data frame with the "ds" column that contains the dates for predicting (Annalyn & Kenneth, 2017). The make_future_dataframe function takes a model object and a number of times to estimate and produces the accurate data frame. It automatically adds the historical dates so that the model fit is visualised as shown below.

Part of the Output;

Table 4.9 An extract of the days to be forecasted in future

	Ds
417	2022-12-31
782	2023-12-31
1147	2024-12-31
1512	2025-12-31
1877	2026-12-31

Peter (2017) indicated that the predict technique is applied to each row in the future, and the estimated value is termed *yhat*. If the historical dates were entered, it would give a sample balance. The prediction component in this case is the latest data frame with an *yhat* column and forecast, as well as the constituent's column and periods of risk.

Table 4.10 An extract of forecasts together with their upper and lower bounds

	ds	yhat	yhat_lower	yhat_upper
417	**2022-12-31**	300609830	133368577	457442300
782	**2023-12-31**	318691910	164604572	467434872
1147	**2024-12-31**	332278231	181771680	496555116
1512	**2025-12-31**	341364538	185905142	490201056
1877	**2026-12-31**	345954449	195607553	497193624

4.8 Forecasting Foreign Direct Investment Prediction Model

Figure 4.9 A plot of forecasts together with original data series-FDI

GAM obtained (line coloured in blue), and the model is doing precisely well and closely following well-makers (black dots) with potential to "skip" points of shocks. Gray shading indicates a period of confidence interval, which often extends as it magnifies the natural forecast horizons in all-time series models (Ruey, 2010).

Yearly Trend Seasonality of the Time Series

Other considerations emerge as a result of decomposing a series into its constituents, as indicated;

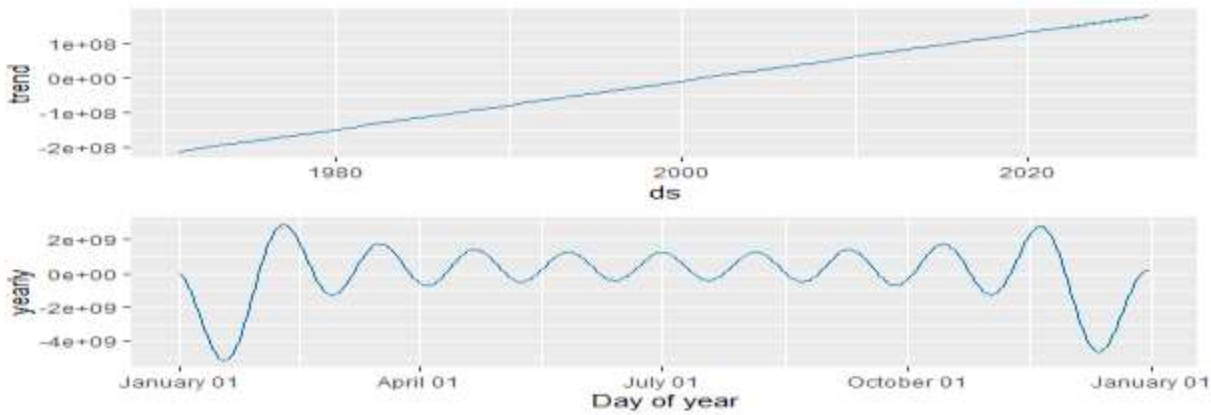


Figure 4.10 Trend and Yearly seasonality of FDI inflows

FDI data shows an increase in trend from 1970 to 2020, despite any other economic changes that took place during that period. It can also be noted that inflows are pretty much lower in February than in January. This is due to changes in wages rates, labour skills, and political stability/ property rights. So this slump acts as a signal to the countries' policymakers to draft out policies that attract the investors there by promoting economic growth and development. However, from March to October, the seasonality trend fluctuated evenly.

4.9 Model performance

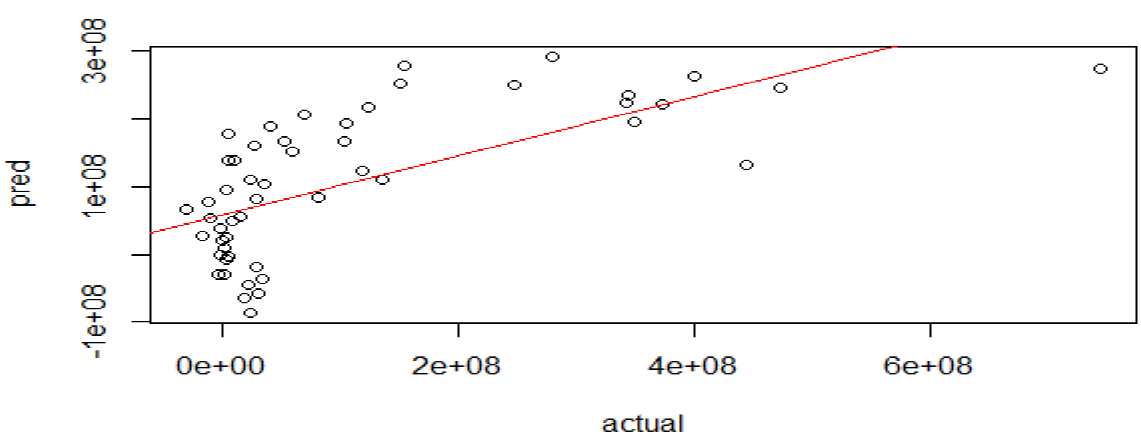


Figure 4.11 Trend and Yearly seasonality of FDI

From the above plot, it can be noted that the model performed well, although at some point it underestimated or overestimated the values. However, other values are closely packed on the red smooth line fitted between the actual values and the predicted values as highlighted by (Annalyn & Kenneth, 2017). This clearly shows a series of correlations of the error term of predicted values on actual values from one period to another.

Table 4.11 Summary actual vs predicted values of FDI inflows

<table>
<tr><td>Residuals:</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Min</td><td>1Q</td><td>Median</td><td>3Q</td><td>Max</td><td></td></tr>
<tr><td>-154840991</td><td>-60369723</td><td>-273784</td><td>68129486</td><td>152226446</td><td></td></tr>
<tr><td>Coefficients:</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td>Estimate</td><td>Std. Error</td><td>t value</td><td>Pr(>|t|)</td><td></td></tr>
<tr><td>(Intercept)</td><td>5.93E+07</td><td>1.34E+07</td><td>4.418</td><td>5.36E-05</td><td>***</td></tr>
<tr><td>Actual</td><td>4.35E-01</td><td>6.99E-02</td><td>6.216</td><td>1.01E-07</td><td>***</td></tr>
<tr><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td colspan="6">Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1</td></tr>
<tr><td colspan="6">Residual standard error: 81070000 on 50 degrees of freedom</td></tr>
<tr><td colspan="2">Multiple R-squared:</td><td>0.4359,</td><td colspan="2">Adjusted R-squared:</td><td>0.4246</td></tr>
<tr><td colspan="2">F-statistic:</td><td colspan="2">38.64 on 1 and 50 DF,</td><td>p-value:</td><td>1.01E-07</td></tr>
</table>

Hence, since the p-value is too small at 1.009e-07 and the R-squared is too high at 0.4359, this indicates that there is a high confidence that this is statistically significant and hence the model did perform well.

CHAPTER 5: SUMMARY, RECOMMENDATIONS AND CONCLUSION

5.0 Introduction

This chapter summarises study results in accordance with the objectives, explicitly indicating the degree or level to which the objectives have been met. It reaches required judgments on the extent to which the results confirm or differ from empirical findings in the same area of research. This chapter aims at providing suggestions in relation to the research findings. To conclude the chapter, recommendations for ongoing studies that supplement forecasting Foreign Direct Investment to Zimbabwe are provided that were not discussed in this research study.

5.1 Summary of findings from the study

The primary objective of this study was to forecast FDI to Zimbabwe using time series analysis from 1970 to 2021. The study found that more conclusions were drawn from the GAM developed and the Foreign Direct Investment GAM revealed that there were a lot of rapid shocks or changes in inflows. As a result, it is worth noting that one of these turning points happened between December 2020 and June 2021, at which time the country suffered a national lockdown and a growing inflation rate due to economic sabotage. This might signal that international investors and FDIs have lost faith in investing in Zimbabwe as a result of the central government's unfavourable economic policies.

GAM, unlike ARIMA models, decomposes sequential data in order to extract a yearly trend using Prophet. As a result of these plots, economic experts and policymakers advise investors and the government on new investment terms that are beneficial to both sides in order to ensure a smooth flow of FDI and thereby promote economic growth in Zimbabwe. Annual patterns, for example (Figure 4.9), show that FDI inflows are lowest in January and highest in June. As a result, the best time to invest is when interest rates are low (January). ARIMA models cannot provide such a detailed analysis in order to answer both policymakers and investors' concerns.

Furthermore, it was discovered that GAM generated more precise results than frequently utilised ARIMA models in predicting the future. Simply put, the MSE obtained from GAM predictions was less than 113742389 for ARIMA (0, 1, 1) model forecasts. Regardless of the fact that perhaps the FDI data under study is not particularly volatile, the MSE from the ARIMA (0, 1, 1) model may imply or hint that there is far more room or a need for further refining and adjusting it.

5.2 Conclusion

The research concluded that GAM surpasses ARIMA models in FDI modelling because it assists investors to address numerous concerns about investing their hard-earned money. Although the ARIMA (0, 1, 1) model outperformed the GAM model in terms of in-sample forecasts, it struggled to provide credible and significant out-of-sample forecasts. It is important to note that predicting economic variables such as FDI is not an easy process because the results of such studies are easily affected by structural breaks in the economy. The results of the estimated models remain relevant as long as there are no structural breaks such as government or disease outbreaks such as COVID19. Therefore, policy makers ought to pay attention to the risk of adjustment in economic operations and maintain the stability and continuity of microeconomic regulation and control in order to prevent the economy from severe fluctuation and adjust to the corresponding target value according to the actual situation.

5.3 Recommendations

The policy implication of this study is that policy makers in Zimbabwe should sustain their effort to expand the tax base, reduce red tape, and strengthen the regulatory framework to investment and investors friendly policies in order to attract the much needed FDI. The political economy of Zimbabwe cannot be overlooked each time economic policy dynamics are discussed (Nyoni, 2018). A stable political environment is recommended because it is a signal of a favourable investment climate. The study can be extended to non-linear models to take care of volatility which may occur in the FDI series.

REFERENCES

Adebiyi, A. A., Adewumi, A. O., & Ayo, C. K. (2014). *Stock Price Prediction Using the ARIMA Model. 2014 UKSim-AMSS 16th International Conference on Computer Modelling and Simulation. Cambridge,UK: UKSim-AMSS*. 105.

Adhikari, R. (2014). *An introductory study on time series modelling and forecasting. Retrieved 03 20, 2015)*. https://doi.org/10.21511/ppm.14(2-2).2016.04

Annalyn, N., & Kenneth, S. (2017). *Time Series Analysis with Generalised Addative Models. Layman Tutorials in Analytics*. 1–5.

Australian Bureuau of statistics. (2008). *Retrieved 05 22, 2017, from www.abs.gov.au.*

Bayai, I., & Nyangara, D. (2013). *An analysis of Determinants of Private Investment in Zimbabwe for the period 2009-2011. International Journal of Economics and Management Sciences. Vol 2, No 6, 2013, pp. 11-42.*

Blog, M. (2008). *Retrieved 05 22, 2015, from wwwjas-style.blogsport.com 2008.*

Clarke, D. G. (1980). *Foreign Companies and International Investment in Zimbabwe. Catholic Institute of International Relations, London.*

Cryer, J. D., & Chan, K. S. (2008). *Time series Analysis with Applications in R. lowa city USA : Springer Texts in Statistics.*

Dawson, C. (2013). *Practical Research Methods: A User-friendly Guide to Mastering Research Techniques and Projects. How To Books, 2002.*

Demaria, F., & Kothari, A. (2017). *The Post-Development Dictionary agenda: paths to the pluriverse, Third World Quarterly, 38:12, 2588-2599, DOI: 10.1080/01436597.2017.1350821*.

George, E., Gwilym, M., & Gregory, C. (. (2008). *Time series analysis.Forecasting and Control. Canada: A John Wiley and sons,inc,publication*.

Gwenhamo, F. (2009). *Foreign Direct Investment in Zimbabwe. The role of Institutional factors, August 17, 2009. ERSA Working Paper 112, Economics Research Southern Africa, Cape.*

Hansen, H., & Rand, J. (2006). *On the Causal Links between FDI and Growth in Developing Countries. World Economy, 29, 21-41. https://doi.org/10.1111/j.1467-9701.2006.00756.x*.

Henry, A. S., Elijah, E. E., Gwani, A. A., & Simon, J. (2019). Time Series ARIMA Model for Predicting Nigeria Net Foreign Direct Investment (FDI). *World Scientific News, 128*(May), 348–362.

Idowu, A. (2021). *Econometric Modelling and Forecasting Foreign Direct Investment Inflows in Nigeria : ARIMA Model Approach*. (107466).

Kothari, C. R. (2017). *Research Methodology Methods and Techniques. 2 nd Edition, New Age International Publishers, New Delhi.*

Makridakis, Hyndman, & Wheelwright. (2007). *Forecasting Methods and Applications. Retrieved 03 09, 2015*.

Moyo, T. (2013). *The impact of Foreign Direct Investment on Economic Growth. The case of Zimbabwe (2009-2012). International Journal of Economics, Finance and Management Sciences. 323-329.*

Nieswiadomy, M. (2010). LSAT® Scores of economics majors: The 2008-9 class update. *Journal of Economic Education*, *41*(3), 331–333. https://doi.org/10.1080/00220485.2010.486739

Nyoni, T. (2018). Box-Jenkins ARIMA approach to predicting net FDI inflows in Zimbabwe. *9th European Conference on Machine Learning Poster Papers*, (87737), 128–137. Retrieved from https://ideas.repec.org/p/pra/mprapa/87737.html

Nyoni, T., & Muchingami, L. (2018). *Foreign Direct Investment (FDI) dynamics in India : what do ARIMA models tell us ?* (93986).

Pannerselvam, R. (2005). *Research Methodology. New Delhi: Prentice hall of India*.

Peter, L. (2017). *Doing magic and analyzing seasonal time series with GAM (Generalized Additive Model) in R.*

Ramasubramanian. (2015). *Time series Analysis. New Delhi: I.A.S.R.I library avenue. This is the Pre-Published Version. (n.d.).*

Robert, & Lipsey, E. (2015). *The national Bureau of Economic Research. Retrieved from Foreign Direct Investment and the Operations of Multinational firms:History and Data: www.nber.org/papers w866*.

Romer, P. (1986). *'Increasing returns and long run economic growth', Journal of Political Economy*. *94*(5), 1002–1037.

Ruey, S. (2010). *Analysis of Financial Time Series Second Edition*.

Taylor, & Letham, B. (2018). *prophet: Automatic Forecasting Procedure, 2019. URL https://facebook.github.io/prophet/*.

Todaro, M., & Smith, S. (2011). *Economic Development. New York: Addison Wesley*.

Todaro, & Smith, S. (2003). *Economic Development,PearsonEducationPte. Ltd., Singapore.* https://doi.org/10.1002/9781119196754.ch2

Velde, D. W. (2006). *ODI Annual Reports. Retrieved from www.od.org/publications/594-foreign-direct-investment-development-historical perspective*.

APPENDICES

Code 1.1

```
###Data was imported to R-Studio from Excell in CSV form using the following command

data <- read.csv(file.choose(), head = T)

View(data)

library(forecast)

library(tseries)

library(ggplot2)

library(MASS)

###Checking whether the data is time series or not

is.ts(data)

###Converting the data to timeseries

DataFrame <- ts(data$FDI_inflows, start = 1970, end = 2020, frequency = 1)

###Checks for missing data

sum(is.na(DataFrame))

###Gives the summary data like mean, median, quartiles, min and max

summary(DataFrame)

###Decomposes the time series into trend, seasonality and random or irregularity components

tsdata=ts(DataFrame, frequency=1)

ddata=decompose(tsdata, "multiplicative")

plot(ddata)

plot(ddata$trend, xlab="Years", ylab="FDI_inflows")

plot(ddata$seasonal, xlab="Years", ylab=" FDI_inflows ")

plot(ddata$random, xlab="Year", ylab=" FDI_inflows ")

###Fits a regression line on the data to identify the trend

plot(DataFrame, xlab="Year", ylab="FDI_inflows")

abline(reg=lm(DataFrame~time(DataFrame)))

###KPSS test for trend stationarity
```

```
kpss.test(DataFrame, null = "Trend", lshort=F)

###ADF test for stationarity

adf.test(DataFrame)

###Fits and gives the optimal model and their estimated parameters

mymodel=auto.arima(DataFrame)

mymodel

summary(mymodel)

plot.ts(mymodel$residuals, xlab="Year", ylab="Model Residuals")

hist(mymodel$residuals, xlab = "Residuals", ylab = "Frequency", main="Histogram of Residuals")

qqnorm(mymodel$residuals)

qqline(mymodel$residuals)

acf(ts(mymodel$residuals), main='Autocorrelation Function of Residuals')

pacf(ts(mymodel$residuals), main='Partial Autocorrelation Function of Residuals')

###Plots the forecasted data

myforecast=forecast(mymodel, level=c(95), h=1*5)

plot(myforecast,xlab="Years", ylab="FDI_inflows")

myforecast

###Model validation

Box.test(mymodel$resid, lag=5, type ="Ljung-Box")
```

Code 1.2

```
library(prophet)

library(dplyr)

library(tidyr)

library(tidyverse)

library(timeDate)

library(lubridate)

library(Metrics)

tsc <- read.csv("C:/Users/MELUSI/Desktop/Dissertation data.csv")
```

```
colnames(tsc)=c('Date','FDI_inflows')

head(tsc)

tsc$Date=ymd(tsc$Date)

tsc$FDI_inflows=as.numeric(tsc$FDI_inflows)

str(tsc)

qplot(Date,FDI_inflows,data=tsc,main='Foreign Direct Investment Inflows in Zimbabwe')

ds=tsc$Date

y=tsc$FDI_inflows

df=data.frame(ds,y)

head(df)

model1=prophet(df)

future1=make_future_dataframe(model1,periods=1*365)

tail(future1)

forecast1=predict(model1,future1)

tail(forecast1[c('ds','yhat','yhat_lower','yhat_upper')])

dyplot.prophet(model1,forecast1,main='Forecasting the GAM Model from Dec 2021 to Dec 2025',ylab='FDI_inflows',xlab='Time')

prophet_plot_components(model1,forecast1)

pred=forecast1$yhat[1:52]

actual=model1$history$y

plot(actual,pred)

abline(lm(pred~actual),col='red')

summary(lm(pred~actual))
```

Sibanda Methembeni Turnitin %

ORIGINALITY REPORT

15%	14%	5%	6%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	mpra.ub.uni-muenchen.de Internet Source	3%
2	elibrary.buse.ac.zw:8080 Internet Source	1%
3	github.com Internet Source	1%
4	ir.buse.ac.zw:8080 Internet Source	1%
5	atsa-es.github.io Internet Source	1%
6	www.worldscientificnews.com Internet Source	1%
7	liboasis.buse.ac.zw:8080 Internet Source	1%
8	Submitted to RMIT University Student Paper	1%
9	Submitted to University of East Anglia Student Paper	1%