UNIVERSITAT DE BARCELONA

SutteARIMA is a New Approach to Forecast Economics, Business, and Actuarial Data

Ansari Saleh Ahmar

Aquesta tesi doctoral està subjecta a la llicència ***Reconeixement- CompartIgual 4.0. Espanya de Creative Commons***.

Esta tesis doctoral está sujeta a la licencia ***Reconocimiento - CompartirIgual 4.0. España de Creative Commons.***

This doctoral thesis is licensed under the ***Creative Commons Attribution-ShareAlike 4.0. Spain License.***

PhD in Business

SutteARIMA is a New Approach to Forecast Economics, Business, and Actuarial Data

Ansari Saleh Ahmar

UNIVERSITAT DE BARCELONA

2022

PhD in Business | Ansari Saleh Ahmar

PhD in Business

Thesis title:

SutteARIMA is a New Approach to Forecast Economics, Business, and Actuarial Data

PhD student:

Ansari Saleh Ahmar

Advisor:

Eva Boj del Val

Date:

January 2022

LIBERTAS PERFVNDET · OMNIA LVCE

UNIVERSITAT DE BARCELONA

Dedicated to my family

ACKNOWLEDGEMENTS

Praise be to Allah SWT, Greetings and respect for the Prophet Muhammad SAW who has given the strength and capacity of Faith and Pious to me so that I can complete my PhD. I would like to express my sincere thanks and highest appreciation to my Supervisor/Advisor (Dr. Eva Boj del Val) for his continued guidance and encouragement, as well as his intellectual input, recommendations and valuable time so that I can complete my PhD. I found his support invaluable since she became my PhD at the start of my research.

I also thank Business School of Universitat de Barcelona staff for their support throughout the school years. I would like to specifically: Jordi Roca and Joana Rafecas for their support during the school days. I also thank several other professors: Professor Dr. M. A. El Safty (Taif University, Saudi Arabia), Dr. Pawan Kumar Singh (Thapar Institute of Engineering and Technology, India), and Dr. Nguyen Van Thanh (Van Lang University, Vietnam) for their useful suggestion, discussion, cooperation, and collaboration in writing articles related to SutteARIMA. Especially for Professor Abdul Rahman and Dr. Rusli who is my "adoptive father" in the world of education for his support and advice during lectures.

Many thanks to Universitat de Barcelona (UB). By studying at UB, I gained weekly research seminars and short courses which were very useful in the Ph.D. process. I also thank the academic commission of Business School of UB for its valuable and useful input. I also thank to Ministry of Education, Culture, Research, and Technology of Republic of Indonesia; Rector of Universitas Negeri Makassar; and Dean of Faculty of Mathematics and Natural Sciences of Universitas Negeri Makassar for their support in continuing and completing my studies at UB.

In particular, I would like to express my sincere appreciation to my parents, in-laws, uncles and aunts as well as my siblings for the support and moral encouragement given to me. I also express my sincere appreciation to my wife (Andi Nurani Mangkawani Arifin), my childrens (Alfatih Abqary Ahmar & Almeera Azzahrah Mecca Ahmar) for their time, patience, understanding, sacrifice, and full support towards the completion of this study. You all are my motivation to complete this research. Finally, I would like to thank all my my parents (Abd. Azis M and Sitti Hamsinah R.), mother-in-law (Iswati Djuaeni), and all my family and my friends at PPI Spanyol especially: Mr. Mahir, Mr. Handoyo, Mrs. Caesar, & Mr. Asdi, and I would like thanks to Embassy of the Republic of Indonesia in Madrid for their best wishes and moral support for my Ph.D.

ABSTRACT

SutteARIMA is a short-term forecasting method developed by combining the α-Sutte Indicator and autoregressive integrated moving average (ARIMA) forecasting methods. SutteARIMA also uses three past time series data in forecasting. In this study, it will be discussed in three specific sections: (1) predict the short-term of confirmed cases of covid-19 and IBEX in Spain; (2) novel forecasting method (SutteARIMA) and its application in predicting Infant Mortality Rate data in Indonesia; and (3) proposes a new model, namely the SutteARIMA model, combining the α-Sutte Indicator and ARIMA methods to forecast economic and finance data regardless of whether the data is linear or non-linear (National Currency to US Dollar Spot Exchange Rate for Indonesia and Consumer Price data Index: All Items for Indonesia). Based on this results, we can conclude that the SutteARIMA method is more accurate than other models, based on mean squared error (MSE) and mean absolute percentage error (MAPE). In addition, from the results of research and test results on the generated artificial data, it is also found that SutteARIMA is suitable for Non Trend, Trend, Non Seasonal, Seasonal data, and their combination.

TABLE OF CONTENTS

Acknowledgements v
Abstract vii
Table of Contents ix
List of Abbreviations xi

Chapter 1 Introduction 1
1.1. Motivation 1
1.2. Objectives 7

Chapter 2 Short-term forecasting method, a case: COVID-19 and stock market in Spain 15
2.1. Introduction 16
2.2. Literature 17
2.3. Methods 21
2.4. Results and Discussion 22
2.5. Conclusion and Further Research 26

Chapter 3 SutteARIMA: A Novel Method for Forecasting the Infant Mortality Rate in Indonesia 23
3.1. Introduction 24
3.2. Literature 26
3.3. Methodology and Data 33
3.4. Results and Discussion 34
3.5. Conclusion and Impact 41

Chapter 4 SutteARIMA: A Novel Forecasting Method for Financial and Economic Data and its Comparison with Other Forecasting Methods 43
4.1. Introduction 44
4.2. SutteARIMA approach 46
4.3. Methodology and data 60
4.4. Empirical Examples 63
4.5. Conclusion, Policy implication, and further research 73

Chapter 5 Conclusions 77
5.1. Main Conclusions 78
5.2. Limitations and Future Research 78

References 81

LIST OF ABBREVIATIONS

ANN	Artificial Neural Network
AR	Autoregressive
ARFIMA	Autoregressive Fractionally Integrated Moving Average
ARIMA	Autoregressive Integrated Moving Average
ARIMAX	Autoregressive Integrated Moving Average Exogenous
ARMA	Autoregressive Moving Average
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
MA	Moving Average
MAPE	Mean Absolute Percentage Error
MSE	Mean Squared Error
NNETAR	Neural Network Time Series Autoregressive
RMSE	Root Mean Squared Error
SVR	Support Vector Regression
VIX	Volatility Index

Chapter 1
Introduction

1.1. Motivation

The International Monetary Fund (IMF) in IMF (2018) explained that in 2017, world growth has strengthened to 3.8 percent and has significantly increased in a global trade. Global growth was projected to rise to 3.9 percent in 2018 and 2019 before 3.8 percent in 2017. This growth was driven by an increase in the projected growth in developing markets and developing economies as well as rapid growth in developed countries. IMF also expected growth for 2018 and 2019 to increase by 0.2 percent annually compared to World Economic Outlook (WEO) in October 2017. In addition, IMF explained that the increase was also driven by the recovery in investment in developed countries, strong economic growth in developing countries in Asia, progress in developing countries of Europe, and signs of recovery in some commodity exporters. Furthermore, this growth was also supported by a strong impetus, the good market sentiment, the accommodative financial conditions, as well as domestic and international impact of the expansionary fiscal policy in the United States. Recovery in some commodity prices should allow a gradual increase in commodity exporters.

In addition to problems in the economic sector as mentioned above, the insurance sector also have uncertainty and sometimes there is an increased demand on the insurance risk. Emamgholipour, Arab, & Mohajerzadeh (2017) explained that if inflation increased by 1 percent, the demand for life insurance will decreases by 0.77 percent, when GDP increases by 1 percent, the demand for life insurance increases by 0.92 percent. Also, the increase in population has a significant positive effect on the demand for life insurance.

Furthermore, Haiss & Sümegi (2008) indicate that there is a significant relationship between insurance and economic growth, but also stress the importance of interest rates and the level of economic development for the insurance sector.

With respect to the economic and insurance uncertainty, this should be anticipated with the approach in predicting these uncertainties, one of them by using data forecasting approach. Forecasting methods are often used to predict stock returns, population, as well as other economic, insurance or reinsurance data, among others. Forecasting has an important role in all sectors, especially in the economic sector, and can help people in decision making processes. Forecasting is often done by various parties while making decisions, for example in economic policy, share, or otherwise. Decision making is usually done in order to foresee events that will occur in the future by using past data.

Uncertainty in the economic sector impacts the investment performance of the sector. If these impacts can be anticipated, they will not produce a negative impact, but if they cannot be anticipated, they will result on a flagging world economy. Time series data changes from time to time and sometimes in an abruptly manner. To view these changes from time to time, estimates of the data needs to be done.

Furthermore, since late 2019 until the present moment, the World is overwhelmed by the pandemic outbreak known as Corona Virus Disease (COVID-19). COVID-19 has the potential to create devastating all sectors of life in terms of social, economic, industrial, and religious crises. COVID-19 is almost and even disrupting economic activity in all countries affected by COVID-19. Economic activities are paralyzed due to social restrictions that include social distancing and the implementation of quarantine/lockdown of certain areas. A lockdown policy refers to a type of quarantine in which citizens in a certain region are prohibited from going in and out of the territory without official permission from authorities to prevent the spread of COVID-19. The community is also regulated in such a way as not to roam around and gather in public places and/or stay at home.

Regional quarantine will have an impact on the economy, the economy will experience a decline due to a lack and/or cessation of production activities and only relies on basic consumption of the community. Stop of all production activities means that there will be closure of industries, and this will result in the laid off on a large scale, because the industry is not able to pay employee salaries. Furthermore, if the quarantine of this area is carried out for a long time, this will definitely result in the collapse of several industrial companies. A number of key economic sectors can be identified as suffering from a drastic fall in output, including accommodation and food services sector with 144 million workers, the wholesale and retail trade segment with 482 million workers, business services and administration with 57 million workers, and manufacturing sector, which employs 463 million workers, has been hit hard in some segments (United Nations, 2020).

Moreover, the regional quarantine policy will also influence the psychology of the community to prepare and calculate their daily needs, so that this can lead to panic buying. This panic buying will have an impact on the unequal distribution of goods and services and can be used by the "mafia" to hoard goods which results in higher price due to an increase in demand for staple goods and services. The high level of unemployment and the price of an item will have an impact on the high inflation in a country (Dixon, 2020).

In addition, the transportation sector will also experience the effects of the regional quarantine due to the disruption of online/conventional transportation operations schedules, trains and airlines. These transportations will lose passengers due to the community restrictions on their activities. Furthermore, the tourism sector is also severely affected by the quarantine restrictions in this region. Tourism is one of the biggest contributors to the formation of a country's Gross Domestic Product (GDP) (Chakrabortya & Maity, 2020).

All of the above conditions are conditions that cause uncertainty which require an approach that can be used to minimize all such conditions, one way that can be done is by forecasting. By predicting the future conditions, then it will be able to provide "enlightenment" about the decisions to be taken by decision makers.

Forecasting has been studied by various academicians, Torri & Vaupel (2012) examined the data forecasting of life expectancy in an international context, they used Italian mortality data (1872-2006) and the United States (1933-2006), this study used ARIMA method. Furthermore, Nyoni & Nathaniel (2018) investigated the use of Autoregressive Moving Average (ARMA), ARIMA, and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) method to modeling rates of Inflation in Nigeria data, the results of this study concluded that ARMA (1, 0, 2) model is clearly the best optimal model based on the minimum Theil's U forecast evaluation statistic compared to the ARIMA (1, 1, 1) model and the AR (3) – GARCH (1, 1) model and based to their diagnostic tests also indicate that the presented models are stable and hence reliable. Abonazel & Abd-Elftah (2019) study about the forecasting GDP of Egyptian using ARIMA, they use annually data from the World-Bank for the years 1965 to 2016, based this study, they conclude that the appropriate statistical model for GDP of Egyptian is ARIMA (1, 2, 1) and doing this model to forecast the GDP of Egypt for the next ten years. Furthermore, Mahmudah (2017) also predicting unemployment rate in Indonesia using ARIMA, and they get result that the best model for this predicting is ARIMA (0,2,1) with average of the residuals is close to zero that indicate that the result is good for forecasting analysis. Not only economics data, we also can forecast mortality and deats data. Rajia, Sabiruzzaman, Islam, Hossain, & Lestrel (2019) study about trends and future of maternal and child health in Bangladesh using ARIMA; Mishra, Sahanaa, & Manikandan (2019) study about forecasting Indian infant mortality rate: An application of autoregressive integrated moving average model; Patowary, Dutta, Barman, & Gadde (2018) study about forecasting accidental deaths in India using ARIMA Model; and Lin, Chen, Chen, Wu, & Lin (2015) study about predicting injury mortality in Xiamen, China using ARIMA.

On the other hand, the ARIMA model not always suitable for predicting data. For example, the forecasting results carried out by the Wang (2011) with the forecasting Taiwan export with fuzzy time series model and ARIMA model show that fuzzy time series models can be utilized to predict export values accurately, outperforming the ARIMA model. Wang (2011) said that the ARIMA model can compensate and balance the value of exports which fluctuate over time in a long period but the shorter the testing period, the more effective the fuzzy time series model is than the ARIMA model. This discrepancy is influenced by Seasonal, Trend, and Cyclic factors. To address this problem, it will be conducted a study of the development of new methods of forecasting by taking into account the seasonal, trend, and cyclic factors. Forecasting method as has been discussed previously by Box and Jenkins that developed a peralama method of ARIMA that accommodated the seasonal, trend, and cyclic factors, but the problems found by using this method is the inaccurate forecasting value obtained thus developed another method to keep using ARIMA forecasting method principle but do modification coefficients. This new forecasting method called SutteARIMA (SF). SutteARIMA also accommodates factors that can affect forecasting results as contained in ARIMA method. The advantages offered from this SutteARIMA method as in the level of accuracy in predicting the data more suitable when compared with ARIMA methods or other methods. This is supported by the results of study conducted by researchers regarding the comparison of results of SF forecasting with other forecasting methods: ARIMA, Fuzzy Time Series, and other methods. Comparison of this forecasting results was obtained by using the comparative value of MSE, root mean squared error (RMSE), and Statistics. Regarding to the forecasting results, it was obtained that Sutte Indicator method is more suitable in forecasting economics/mortality data when compared to other methods. Sutte Indicator was discussed by Ahmar, Rahman, Arifin, & Ahmar (2017) about the prediction of movement stock; Kurniasih, Ahmar, Hidayat, Agustin,

& Rizal (2018) about forecasting infant mortality rate for China using α-Sutte Indicator, ARIMA, and Holt-Winters; Sutiksno, Ahmar, Kurniasih, Susanto, & Leiwakabessy (2018) about forecasting historical data of Bitcoin using ARIMA and α-Sutte Indicator; and Ahmar, Rahman, & Mulbar (2018) about α-Sutte Indicator, the new method for time series forecasting. SutteARIMA is the development of method of Sutte Indicator, α-Sutte Indicator and combine with ARIMA method.

In addition to this method, another method is also carried out by various researchers, namely the incorporation of several forecasting methods. Based on the literature, several researchers (Clemen, 1989; Hibon & Evgeniou, 2005; Winkler & Makridakis, 1983) have empirically shown the effectiveness of combination based forecast. (Becker & Clements, 2008) did the forecasting to the S&P 500 data with the model confidence set (MCS) method developed by (Hansen, Lunde, & Nason, 2003), this study combines combinations of ARMA+Autoregressive Fractionally Integrated Moving (ARFIMA) and ARMA+ARFIMA+SVR+Volatility Index (VIX) and compares forecast results based on MSE and QLIKE values. (Aiolfi & Timmermann, 2006) said that:

> "We proposed a set of new combination strategies that first sort models into either quartiles or clusters on the basis of the distribution of past forecasting performance across models, pool forecasts within each cluster and then estimate optimal combination weights and shrink these towards equal weights. This combination scheme makes use of many of the techniques proposed in the literature for **improving forecast combinations** such as trimming, pooling, optimal weighting and shrinkage estimation. We find evidence in our data that these conditional combination strategies lead to better overall forecasting performance than simpler strategies in common use such as using the previous best model or simply averaging across all forecasting models or a small subset of these."

Furthermore, (Pauwels & Vasnev, 2014) did the forecasting to the US business cycle index and concluded that the forecast accuracy improves when combining the probability forecasts of both the coincident indicators model and the yield curve model, compared to each model's own forecasting performance.

Based on the results of a preliminary study (Ahmar, 2019), which conducted a simulation trial for data containing Non Trend, Trend, Non Seasonal, Seasonal and compared several forecasting methods, namely α-Sutte, ARIMA, Holt-Winter, Neural network time series forecasts (NNETAR), Robust, Theta, and SutteARIMA (obtained from R Package e.g. SutteForecastR (Ahmar, 2017)). The data generated were 200 data which were divided into 193 data trains and 7 data for forecast. Based on the results of this simulation, it was obtained that:

- NNETAR & Theta are more suitable for Non Trend, Non Seasonal/Cyclic data,
- Robust and Theta are more suitable for Up Trend, Non Seasonal/Cyclic data,
- SutteARIMA is suitable for Down Trend, Non Seasonal/Cyclic data,
- α-Sutte is suitable for Up & Down Trend, Non Seasonal/Cyclic data,
- SutteARIMA is suitable for Up & Down Trend, Seasonal/Cyclic data,
- Holt-Winter and NNETAR are suitable for Up Trend, Non Seasonal/Cyclic data,
- SutteARIMA is suitable for Trend, Seasonal/Cyclic data,
- Holt-Winter is suitable for Trend, Non Seasonal/Cyclic data,
- Holt-Winter is suitable for Non Trend, Non Seasonal/Cyclic data,
- ARIMA and SutteARIMA are suitable for Non Trend, Seasonal/Cyclic data,
- SutteARIMA is suitable for Up & Down Trend, Seasonal/Cyclic data.

1.2. Objectives

1) General objectives

The main objective of this study was to develop a new forecasting method called SutteARIMA method which was developed by using a combination and/or weaknesses of some forecasting methods that already exist. One that is usually required to perform forecasting is data to be stationary, the developed SutteARIMA method will accommodate the assumption that the data is not necessarily stationary.

2) Specifics objectives

To see more details about the purpose of this study, it will be outlined into three specific objectives as follows:

a) Development of new forecasting method (SutteARIMA) in the sector of finance.

To develop a new method (formula) there should be a review process concerning methods of forecasting that already exist. One method that can be done is a review of research methods by using review. In this review, it will be obtained information on the various types of forecasting methods used. From the results of this review, it can be found the weaknesses and constraints of these methods so that, it can be minimized on the SutteARIMA method.

SutteARIMA method which is the first objective of this study has been successfully developed and has been published in the *Science of the Total Environmental Journal*, Volume 729, with the title of the article: *SutteARIMA: Short-term forecasting method, a case: COVID-19 and stock market in Spain* (Ahmar & Boj, 2020).

The purpose of our research is to predict the short-term of confirmed cases of covid-19 and IBEX in Spain by using SutteARIMA method. Covid-19 Spanish confirmed data obtained from Worldometer and Spain Stock Market data (IBEX 35) data obtained from Yahoo Finance. Data starts from 12 February 2020 – 09 April 2020 (the date on Covid-19 was detected in Spain). The data from 12 February 2020 – 02 April 2020 using to fitting with data from 03 April – 09 April 2020. Based on the fitting data, we can do short forecast for 3 future period (10 April – 12 April 2020 for Covid-19 and 14 April – 16 April 2020 for IBEX). In this study, the SutteARIMA method will be used. For the evaluation of the forecasting methods we applied forecasting accuracy measure mean absolute percentage error, MAPE. Based on the results of ARIMA and SutteARIMA forecasting methods, we conclude that the SutteARIMA method is most suitable than ARIMA to calculate the daily forecasts of confirmed cases of Covid-19 and IBEX in Spain. The MAPE value of 0.1905 (smaller than 0.04 compared to MAPE value of ARIMA) for confirmed cases of Covid-19 in Spain and 0,0202 for IBEX stock. At the end of the analysis, using the SutteARIMA method, we calculate daily forecasts of confirmed cases of Covid-19 in Spain from 10 April 2020 until 12 April 2020 and Spain Stock Market from 14 April until 16 April 2020.

This paper discussed about the process of developing the formula, and the formula has been obtained from the SutteARIMA method which is then implemented in predicting the case of COVID-19 and the Stock Market in Spain. From the results, it was obtained that the SutteARIMA method has a better level of accuracy compared to the ARIMA method seen from the mean absolute percentage error, MAPE, value.

b) Forecasting of financial and actuarial data by using SutteARIMA method

In previous studies, it has been discussed about the review that will provide recommendations for the development of SutteARIMA method. SutteARIMA method then conducted continuous process of testing and using the system of trial and error to see the strengths and weaknesses of SutteARIMA. After the ideal method of SutteARIMA is obtained, then that method will be tested in the forecasting process using trend data and linear data.

c) Comparative study of financial data forecasting results

After the ideal method of SutteARIMA is obtained and have passed the testing phase, then the next step is to compare the results of forecasting that is done with SutteARIMA method and other similar methods of forecasting. This comparison is to look at the extent to which SutteARIMA accuracy level when compared with other methods. For this purpose, a forecasting test will be carried out by using various datasets to determine the reliability of the SutteARIMA method.

The first paper of the second and third objective has been published in the *CMC-Computers, Materials & Continua Journal*, Volume 70, Number 3, with the title of the article: *SutteARIMA: A Novel Method for Forecasting the Infant Mortality Rate in Indonesia* (Ahmar et al., 2022).

This study focuses on the novel forecasting method (SutteARIMA) and its application in predicting Infant Mortality Rate data in Indonesia. It undertakes a comparison of the most popular and widely used four forecasting methods: ARIMA, Neural Networks Time Series (NNAR), Holt-Winters, and SutteARIMA. The data used were obtained from the website of the World Bank. The data consisted of the annual infant mortality rate (per 1000 live births) from 1991 to 2019. To determine a suitable and best method for predicting Infant Mortality rate, the forecasting results of these four methods were compared based on the MAPE and MSE. The results of the study showed that the accuracy level of SutteARIMA method (MAPE: 0.83% and MSE: 0.046) in predicting Infant Mortality rate in Indonesia was smaller than the other three forecasting methods, specifically the ARIMA (0.2.2) with a MAPE of 1.21 % and a MSE of 0.146; the NNAR with a MAPE of 7.95% and a MSE of 3.90; and the Holt-Winters with a MAPE of 1.03% and a MSE: of 0.083.

The second paper of second and third objective, in process review in:

with the title of article: *SutteARIMA: A Novel Forecasting Method for Financial and Economic Data and its Comparison with Other Forecasting Methods*.

This paper proposes a new model, namely the SutteARIMA model, combining the α-Sutte Indicator and ARIMA methods to forecast economic and finance data regardless of whether the data is linear or non-linear. The proposed model has the following advantages: it does not pay attention to linear-nonlinear data, the forecasting accuracy is more stable, and the calculation rate is faster. To evaluate the performance, SutteARIMA is compared with conventional models provided (linear and non-linear) namely ARIMA, Holt-Winters, Neural Network, Robust, and Theta models. The results of the study showed that the SutteARIMA method is more accurate than other models, based on MSE: 66474.88 / MAPE: 1.33% on National Currency to US Dollar Spot Exchange Rate for Indonesia and MSE: 0.0493 / MAPE: 0.1594% on Consumer Price data Index: All Items for Indonesia. In addition, from the results of research and test results on the generated artificial data, it is also found that SutteARIMA is suitable for Non Trend, Trend, Non Seasonal, Seasonal data, and their combination.

Chapter 2
SutteARIMA: Short-term forecasting method, a case: COVID-19 and stock market in Spain

2.1. Introduction

The International Monetary Fund (IMF) in IMF (2018) explained that in 2017, world growth has strengthened to 3.8 percent and has significantly increased in a global trade. Global growth was projected to rise to 3.9 percent in 2018 and 2019 before 3.8 percent in 2017. This growth was driven by an increase in the projected growth in developing markets and developing economies as well as rapid growth in developed countries. IMF also expected growth for 2018 and 2019 to increase by 0.2 percent annually compared to World Economic Outlook (WEO) in October 2017. In addition, IMF explained that the increase was also driven by the recovery in investment in developed countries, strong economic growth in developing countries in Asia, progress in developing countries of Europe, and signs of recovery in some commodity exporters. Furthermore, this growth was also supported by a strong impetus, the good market sentiment, the accommodative financial conditions, as well as domestic and international impact of the expansionary fiscal policy in the United States. Recovery in some commodity prices should allow a gradual increase in commodity exporters.

Today, the world is shocked by the epidemic called COVID-19. COVID-19 is a contagious and deadly disease that currently exists in the world by WHO. COVID-19 was first reported in Wuhan, Hubei Province, China in December 2019. COVID-19 is an infectious disease caused by a new coronavirus (SARS-CoV-2) discovered in China (Yang et al., 2020). Based on WHO (2020) data, as of 6 April 2020, there were 1210956 confirmed cases and 67594 confirmed deaths. In Spain, COVID-19 cases began to be detected on 12 February 2020. The highest addition of COVID-19 cases occurred on 26 March 2020, as many as 8271 cases (Worldometer, 2020). Based on data presented by Worldometer on 8 April 2020, the number of confirmed cases of COVID-19 in Spain was 148,220 people with 14,792 deaths, and 48,021 people recovered and was the second highest country in the world with confirmed cases of COVID-19 (Worldometer, 2020).

To anticipate the many confirmed cases of COVID-19, Spain began lockdown on 14 March 2020 (France24, 2020), this lockdown also resulted in all restaurants, bars, hotels, schools and universities all being closed and of course this will have an impact on the economy of the Spanish country especially Spain Market Index (IBEX 35) which experienced a decline of up to 14% at the closing of shares (McMurtry, 2020).

To see more about the impact of lockdown and COVID-19, it is necessary to forecast the data. Time series data changes from time to time and sometimes in an abruptly manner. To view these changes from time to time, estimates of the data need to be done. Forecasting or predictions related to COVID-19 have been studied by various researchers: (Fanelli and Piazza, 2020) studying the forecasting of the spread of COVID-19 in China, Italy, and France using the SIRD model, (Roosa et al., 2020) studying about COVID-19 real-time forecast in China with generalized logistic growth model (GLM), (Benvenuto et al., 2020) examines the forecast of COVID-19 using ARIMA, and (Koczkodaj et al., 2020) predicts COVID-19 outside of China by using a simple heuristic (exponential curve).

2.2. Literature

1) Autoregressive Integrate Moving Average (ARIMA)

Autoregressive Integrate Moving Average (ARIMA) Model first introduced by George Box and Gwilym Jenkins in 1976. The general, the model of ARIMA written with notation ARIMA (p,d,q), with p represents the order of the autoregressive (AR) process, d represents the differencing, and q states the order of the moving average (MA) process.

a) White Noise Process

In forming a time series model, the data must be stationary.

Definition 2.1 Stationary. The time series process Z_t, $t \in \mathbf{Z}$ define stationarity (or weak stationarity) as follows (Brockwell and Davis, 2016; Montgomery et al., 2015) :

(1) the expected value of the time series does not depend on time, $E(Z_t)$ is independent of *t*, where *t* = time.
(2) the autocovariance function defined as $Cov(Y_t, Y_{t+k})$ for any lag *k* is only a function of *k* and not time; that is, $\gamma(Y_t, Y_{t+k})$, or $\gamma(Y_t, Y_{t+k})$ is independent of *t* for each *k*.

Definition 2.2 $\{a_t\}$ process define ***white noise*** with mean 0 and variance σ^2, (Brockwel and Davis, 2006):

$$\{a_t\} \sim WN(0, \sigma^2).$$

If and only if $\{a_t\}$ meets:

$$\gamma(t)=\begin{cases}\sigma^2, & t=0,\\ 0, & t\neq 0.\end{cases} \quad (2.1)$$

Wei (2006) added that white noise process $\{a_t\}$ stationary with autocorrelation function:

$$\rho_k=\begin{cases}1, & k=0\\ 0, & k\neq 0,\end{cases}$$

and partial autocorrelation function:

$$\phi_{kk}=\begin{cases}1, & k=0\\ 0, & k\neq 0.\end{cases}$$

b) Autoregressive Model (AR)

The autoregressive model is a form of regression that links the observations of a particular moment with the values of previous observations at a specific time interval.

The generally, form of autoregressive process the data order p (AR(p)) formulate as (Wei, 1994):

$$\begin{aligned}
Z_t &= \phi_1 Z_{t-1}+\phi_2 Z_{t-2}+\ldots+\phi_p Z_{t-p}+a_t, a_t\sim WN\left(0,\sigma^2\right), \phi_i\in\mathbb{R}, t\in\mathbb{Z}\\
Z_t &= \phi_1 BZ_t+\phi_2 B^2 Z_t+\ldots+\phi_p B^p Z_t+a_t, a_t\sim WN\left(0,\sigma^2\right), \phi_i\in\mathbb{R}, t\in\mathbb{Z} \qquad (2.2)\\
\left(1-\phi_1 B-\phi_2 B^2-\ldots-\phi_p B^p\right)Z_t &= a_t, a_t\sim WN\left(0,\sigma^2\right), \phi_i\in\mathbb{R}, t\in\mathbb{Z}.
\end{aligned}$$

The equation (2.2) can be simplified $\phi_p(B)Z_t=a_t$, with $\phi_p(B)=1-\phi_1 B-\phi_2 B^2-\ldots-\phi_p B^p$.

c) Moving Average Model (MA)

The moving average process is a process that the time series value at time t is influenced by the current error element and may be weighted in the past.

The general form of the process of moving average order q is expressed by MA (q) (Wei, 1994):

$$Z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \theta_q a_{t-q}, a_t \sim WN\left(0,\sigma^2\right), \theta_i \in \mathbb{R}, t \in \mathbb{Z}.$$
$$Z_t = \sum_{i=0}^{q} \phi_i a_{t-i}, \theta_0 = 1, a_t \sim WN\left(0,\sigma^2\right), \phi_i \in \mathbb{R}, t \in \mathbb{Z}. \quad (2.3)$$

or can simplified $z_t = \theta_q\left(B\right) a_t, a_t \sim WN\left(0,\sigma^2\right), \theta_q \in \mathbb{R}, t \in \mathbb{Z}.$

with $\theta_q(B) = (1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q).$

d) Autoregresive Integrated Moving Average or ARIMA (p,d,q)

The (Z_t) process are an autoregressive-moving average or ARMA (p, q) model if it fulfilled (Wei, 1994):

$$\phi_p(B) Z_t = \theta_q(B) a_t, a_t \sim WN\left(0,\sigma^2\right), \phi_p, \theta_q \in \mathbb{R}, t \in \mathbb{Z}. \quad (2.3)$$

with $\phi_p(B) = (1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p)$ (for AR(p))

and $\theta_q(B) = (1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q)$ (for MA(q)).

If there is a differencing, then the ARIMA model becomes as follows:

$$\phi_p(B)\left(1-B\right)^d Z_t = \theta_q(B) a_t, a_t \sim WN\left(0,\sigma^2\right), \phi_p, \theta_q \in \mathbb{R}, t \in \mathbb{Z}.$$

with $\phi_p(B) = (1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p)$ (for AR(p)), $(1-B)^d$ (for differencing non seasonal) and $\theta_q(B) = (1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q$ (for MA(q)).

2) α-Sutte Indicator

α-Sutte Indicator was developed using the principle of the forecasting method of using the previous data (Ahmar et al., 2018). A was also developed using the adopted moving average method. The moving average method is used to predict the trend history of the data. The α-Sutte Indicator uses 4 previous data (Z_{t-1}, Z_{t-2}, Z_{t-3}, and Z_{t-4}) as supporting data for forecasting and making the decision (Ahmar, 2018).

The equation of the α-Sutte Indicator method are as follows (Ahmar, 2018):

$$Z_t = \frac{\gamma\left(\dfrac{\Delta x}{\dfrac{\gamma+\delta}{2}}\right) + \beta\left(\dfrac{\Delta y}{\dfrac{\beta+\gamma}{2}}\right) + \alpha\left(\dfrac{\Delta z}{\dfrac{\alpha+\beta}{2}}\right)}{3}, \quad (2.4)$$

where:

$\delta = Z_{t-4}$,
$\gamma = Z_{t-3}$,
$\beta = Z_{t-2}$,
$\alpha = Z_{t-1}$,
$\Delta x = \gamma - \delta = Z_{t-3} - Z_{t-4}$,
$\Delta y = \beta - \gamma = Z_{t-2} - Z_{t-3}$,
$\Delta z = \alpha - \beta = Z_{t-1} - Z_{t-2}$,
Z_t = data at t time,
Z_{t-k} = data at $(t - k)$ time.

3) SutteARIMA

SutteARIMA is a forecasting method that combines the α-Sutte Indicator with ARIMA. The result of SutteARIMA are the average forecast results from the α-Sutte Indicator and ARIMA.

The equation (2.3), we can describe:

$$\begin{aligned}&(1-\phi_1 B-\phi_2 B^2-\ldots-\phi_p B^p)Z_t=(1-\theta_1 B-\theta_2 B^2-\ldots-\theta_q B^q)a_t,\\&Z_t-\phi_1 BZ_t-\phi_2 B^2 Z_t-\ldots-\phi_p B^p Z_t=a_t-\theta_1 Ba_t-\theta_2 B^2 a_t-\ldots-\theta_q B^q a_t.\end{aligned} \tag{2.5}$$

If equation (2.5) we reduce using backward shift operator $\left(B^p Z_t = Z_{t-p}\right)$:

$$\begin{aligned}&Z_t-\phi_1 Z_{t-1}-\phi_2 Z_{t-2}-\ldots-\phi_p Z_{t-p}=a_t-\theta_1 a_{t-1}-\theta_2 a_{t-2}-\ldots-\theta_q a_{t-q},\\&Z_t=\phi_1 Z_{t-1}+\phi_2 Z_{t-2}+\ldots+\phi_p Z_{t-p}+a_t-\theta_1 a_{t-1}-\theta_2 a_{t-2}-\ldots-\theta_q a_{t-q}.\end{aligned} \tag{2.6}$$

If we define:

$\delta = Z_{t-4}$,
$\gamma = Z_{t-3}$,
$\beta = Z_{t-2,}$
$\alpha = Z_{t-1}$.

The equation (2.6):

$$Z_t=\phi_1\alpha+\phi_2\beta+\phi_3\gamma+\phi_4\delta+\ldots+\phi_p Z_{t-p}+a_t-\theta_1 a_{t-1}-\theta_2 a_{t-2}-\ldots-\theta_q a_{t-q} \tag{2.7}$$

and the equation (2.4) we can simplify:

$$Z_t = \frac{\gamma\left(\frac{\frac{\Delta x}{\gamma+\delta}}{2}\right) + \beta\left(\frac{\frac{\Delta y}{\beta+\gamma}}{2}\right) + \alpha\left(\frac{\frac{\Delta z}{\alpha+\beta}}{2}\right)}{3},$$

$$Z_t = \frac{\frac{\frac{\gamma\Delta x}{\gamma+\delta}}{2} + \frac{\frac{\beta\Delta y}{\beta+\gamma}}{2} + \frac{\frac{\alpha\Delta z}{\alpha+\beta}}{2}}{3},$$

$$Z_t = \frac{\gamma\Delta x}{\frac{3\gamma+3\delta}{2}} + \frac{\beta\Delta y}{\frac{3\beta+3\gamma}{2}} + \frac{\alpha\Delta z}{\frac{3\alpha+3\beta}{2}},$$

$$Z_t = \frac{2\gamma\Delta x}{3\gamma+3\delta} + \frac{2\beta\Delta y}{3\beta+3\gamma} + \frac{2\alpha\Delta z}{3\alpha+3\beta},$$

$$Z_t = \gamma\frac{2\Delta x}{3\gamma+3\delta} + \beta\frac{2\Delta y}{3\beta+3\gamma} + \alpha\frac{2\Delta z}{3\alpha+3\beta}.$$

Let, Equation (2.4) added with Equation (2.7), we find:

$$\begin{aligned} 2Z_t &= \phi_1\alpha + \phi_2\beta + \phi_3\gamma + \phi_4\delta + \ldots + \phi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \theta_q a_{t-q} + \\ &\quad \gamma\frac{2\Delta x}{3\gamma+3\delta} + \beta\frac{2\Delta y}{3\beta+3\gamma} + \alpha\frac{2\Delta z}{3\alpha+3\beta}, \\ Z_t &= \alpha\left(\frac{\phi_1}{2} + \frac{\Delta z}{3\alpha+3\beta}\right) + \beta\left(\frac{\phi_3}{2} + \frac{2\Delta y}{3\beta+3\gamma}\right) + \gamma\left(\frac{\phi_3}{2} + \frac{2\Delta x}{3\gamma+3\delta}\right) + \\ &\quad \frac{\phi_4\delta}{2} + \ldots + \frac{\phi_p Z_{t-p}}{2} + \frac{a_t}{2} - \frac{\theta_1 a_{t-1}}{2} - \frac{\theta_2 a_{t-2}}{2} - \ldots - \frac{\theta_q a_{t-q}}{2}. \end{aligned} \tag{2.8}$$

So, the equation (2.8) is the formula of SutteARIMA.

2.3. Methods

1) Data

COVID-19 Spanish confirmed data obtained from Worldometer and Spain Stock Market data (IBEX 35) data obtained from Yahoo Finance. Data starts from 12 February 2020 – 09 April 2020 (the date on COVID-19 was detected in Spain). The data from 12 February 2020 – 02 April 2020 using to fitting with data from 03 April – 09 April 2020. Based on the fitting data, we can doing short forecast for 3 future period.

2) Statistical Analysis

In making predictions or forecasting, there are several types of methods that can be used, including ARIMA, Holt-Winters, Double Exponential Smoothing, α-Sutte, SutteARIMA, and others. In this study, the SutteARIMA and ARIMA method will be used. Based on preliminary research of (Ahmar, 2019), the SutteARIMA can predicted the Trend data.

The ARIMA method choose because this method is used by several health researchers to monitoring and predicting the development of a disease, for example: (Anokye et al., 2018) using ARIMA to forecast malaria in Kumasi; (Liu et al., 2011) using ARIMA to forecast incidence of hemorrhagic fever with renal syndrome in China; (Zhang et al., 2014) using two decomposition methods (regression and exponential smoothing), autoregressive integrated moving average (ARIMA) and support vector machine (SVM) to forecast epidemiological surveillance data in Mainland China; (Molina et al., 2018) using ARIMA and Autoregressive Integrated Moving Average Exogenous (ARIMAX) to predict bovine trichomoniasis (BT) and bovine genital campylobacteriosis (BGC) prevalence and persistence in La Pampa (Argentina); (Wang et al., 2018) compare ARIMA and GM(1,1) models to predict the hepatitis B in China.

For the evaluation of the forecasting methods we applied two forecasting accuracy measures, mean absolute percentage error, MAPE (Kim and Kim, 2016).

$$\text{MAPE} = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{A_t - F_t}{At}\right| \tag{2.9}$$

where:

A_t = Actual values at data time t.
F_t = Forecast values at data time t.

The results of this forecasting are obtained by using R Software with the forecast and SutteForecastR Package (Ahmar, 2017).

2.4. Results and Discussion

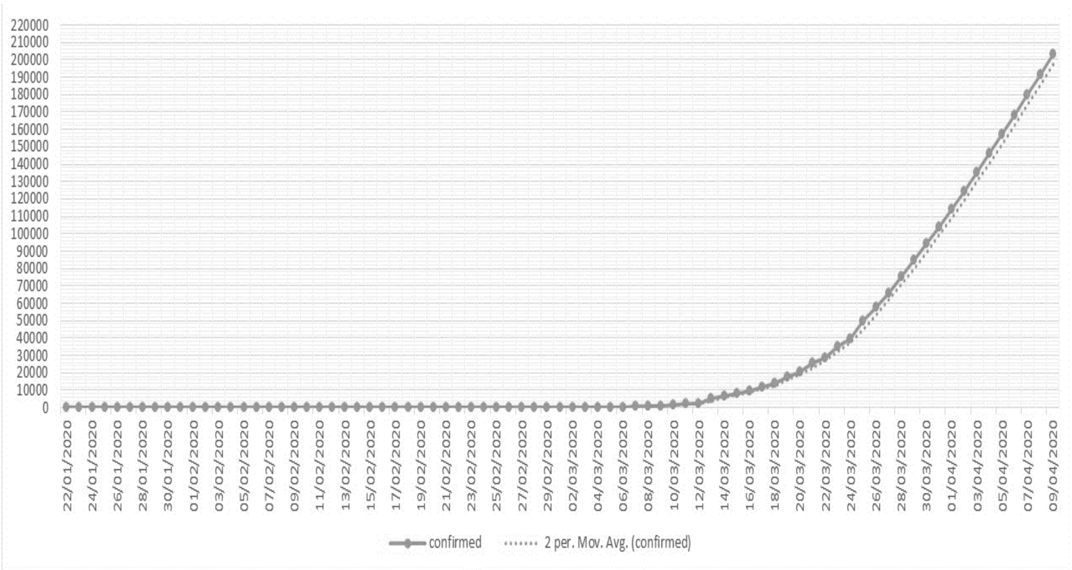
Short-term daily estimates are important for making strategic decisions for the future. In the case of COVID-19, daily forecasting can provide information to decision makers to find a way to prevent COVID-19 from spreading.

Figure 2.1 shows that the confirmed cases of COVID-19 in Spain will continue to grow until this curve is sloped. One of the weaknesses of time series forecasting uses previous data experience as predictive data to be data so that predictions that are suitable for the COVID-19 case are short forecasting for 3-5 future periods. Figure 2.1 also show the addition of confirmed cases of COVID-19 in Spain every day seems to be stabilizing in around 5000 cases.

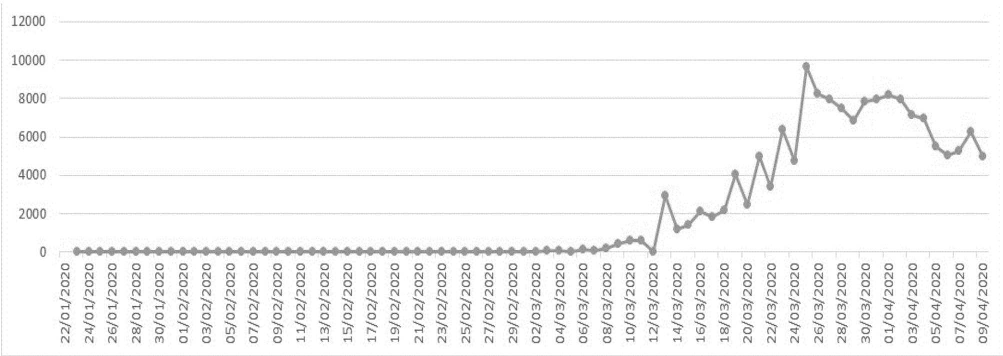
Since COVID-19 is established as a pandemic by WHO and the existence of lockdown or social restrictions will affect the economic development of a country. One thing that is influential is the stock market because with the existence of this pandemic, investors are starting to panic buying, so selling stock has resulted in a drop in stock prices. Moreover, based on WHO data on 9 April 2020, Spain became the second highest country with a confirmed case of COVID-19 in the world.

Figure 2.2 shows that the closing price of the IBEX stock market has decreased from the beginning of COVID-19 in Spain (12 February 2020) and began to stabilize on 24 March 2020 around 6900 per share.

Based on the description, the process of forecasting the data is done using the ARIMA and SutteARIMA methods. The results are presented in table 2.1 for confirmed cases of COVID-19 in Spain and table 2.2 for IBEX Stock.



(a)



(b)

Figure 2.1. (a) Confirmed Cases of COVID-19 in Spain (12 February 2020 – 09 April 2020) (b) Daily New Cases of COVID-19 in Spain (12 February 2020 – 09 April 2020)

Table 2.1. Results of Fitting Confirmed Cases of COVID-19 in Spain

Date	Actual	ARIMA(2,2,1)	APE	SutteARIMA	APE
03/04/2020	119199	120424	0.0103	120425	0.0103
04/04/2020	126168	128740	0.0800	127990	0.0738
05/04/2020	131646	137307	0.1519	135531	0.1370
06/04/2020	136675	145917	0.2241	142134	0.1924
07/04/2020	141942	154699	0.2978	148668	0.2472
08/04/2020	148220	163557	0.3721	155430	0.3040
09/04/2020	153222	172545	0.4475	163200	0.3691
		MAPE	**0.2263**	**MAPE**	**0.1905**

Figure 2.2. Closing Price of IBEX Stock Spain (12 February 2020 – 09 April 2020)

Table 2.1 shows that the SutteARIMA method is most appropriate for predicting the confirmed cases of COVID-19 in Spain with MAPE value of 0.1905 (smaller than 0.04 compared to MAPE value of ARIMA). So, the SutteARIMA method will be used to predict Confirmed Cases of COVID-19 from 10 April to 12 April 2020 (table 2.2).

Table 2.2. Forecast for Confirmed Case of COVID-19 in Spain from 10 April to 12 April 2020

Date	Forecast	Lower 99%	Higher 99%
10/04/2020	158925	157336	160498
11/04/2020	164390	162746	166017
12/04/2020	169969	168269	171651

Based on table 2.3, As the SutteARIMA method is the most appropriate method for this time series, we use it to forecast the IBEX Stock from 10 April 2020 to 12 April 2020 (table 2.4).

Table 2.3. Results of Fitting Data of IBEX Stock

Date	Actual	ARIMA(0,1,0) with drift	APE	SutteARIMA	APE
03/04/2020	6579.4	6692.606	0.017206	6698.858	0.018156
04/04/2020	6574.1	6599.812	0.003102	6557.420	0.003341
05/04/2020	6581.6	6507.018	0.011001	6526.984	0.007967
06/04/2020	6844.3	6414.223	0.025105	6464.475	0.017467
07/04/2020	7002.0	6321.429	0.039209	6627.872	0.007367
08/04/2020	6951.8	6228.635	0.053313	6687.791	0.016474
09/04/2020	7070.6	6135.841	0.067416	6606.707	0.004150
		MAPE	**0.030908**	**MAPE**	**0.010703**

Table 2.4. Forecast for Closing Price of IBEX from 10 April to 12 April 2020

Date	Forecast	Lower 99%	Higher 99%
14/04/2020	7000.61	6930.60	7069.91
15/04/2020	6930.61	6861.30	6999.22
16/04/2020	6860.62	6792.01	6928.54

Based on forecasting results, we can conclude that: SutteARIMA method is the most suitable forecasting method to forecast confirmed cases of COVID-19 in Spain and Closing Price of IBEX. This can be verified by the value of forecasting accuracy measures (MAPE), the best SutteARIMA method for all data.

2.5. Conclusion and Further Research

Forecasting of COVID-19 and IBEX Stock in Spain can give an idea of the policy maker to make decision for future. In fitting data COVID-19 and IBEX Stock in Spain from 3 April 2020 to 9 April 2020, the SutteARIMA method is more suitable than ARIMA method. The confirmed cases of COVID-19 of Spain on 12 April 2020 that is 169969 with interval value 168269-171651 cases and the closing price of IBEX Stock on 16 April 2020 that is 6860.62 with interval value 6792.01 – 6928.54. Based on the forecast, the policy maker can use to make a policy for future. For further research, this method can comparison with other methods, for example with Neural Network.

Chapter 3
SutteARIMA: A Novel Method for Forecasting the Infant Mortality Rate in Indonesia

3.1. Introduction

In this era of globalization and continuous industrial development, every human being wants to get information as fast as possible. Statistics, which is one of the fields of science related to the acquisition of information in several scientific disciplines, has made progress. This advancement usually requires different methods of solving different problems. Statistics has been known for a long time and has even been used in dealing with problems in everyday life such as in the fields of health, economics, social sciences, atmospheric sciences, and other fields. In addition, the development of data mining and big data analysis also requires an understanding of statistics. This is in line with the opinion of Sivarajah et al. (2017) in their presentation depicted in Figure 3.1. Figure 3.1 shows that the types of classification of big data analytical methods, especially in the descriptive analytical, inquisitive and predictive analytical sections require statistical analysis to obtain information. Furthermore, Grover & Mehra (2008) stated that data mining is the application of statistics in the form of exploratory analysis and modeling of data to obtain shapes and trends from large data sets.

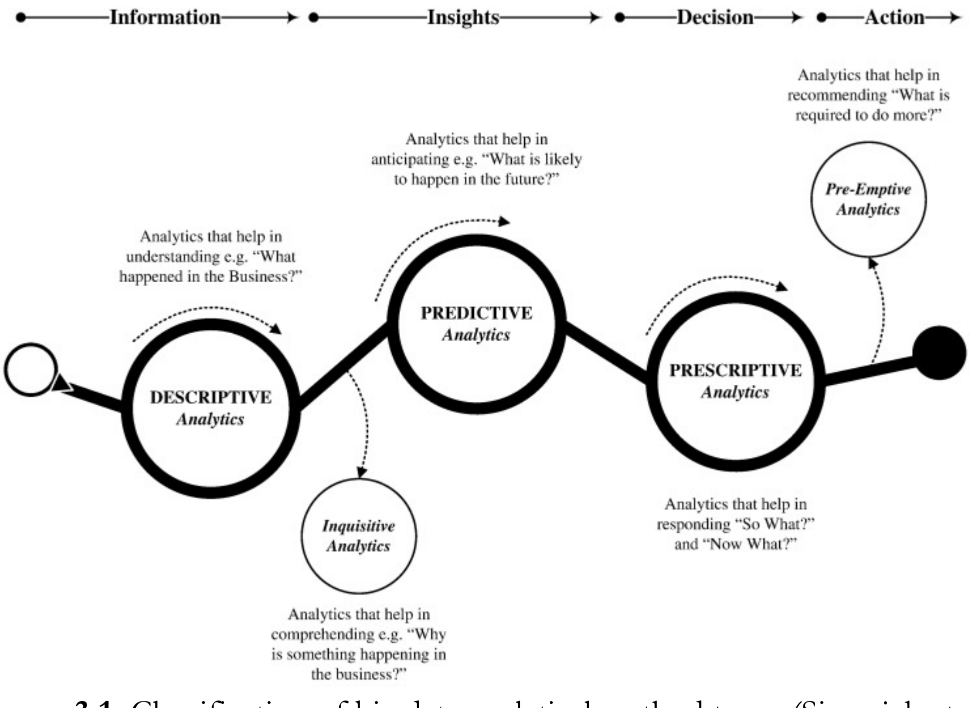


Figure 3.1. Classification of big data analytical method types (Sivarajah et al., 2017)

Statistics are usually used by data analysts to consider possible events that may recur. Therefore, the likelihood of future events is strongly influenced by the frequency and routine of events that have occurred in the past. This is in line with the opinion of Edwards (2019) who states that predictive analysis is data analysis that aims to make predictions about future events based on historical data and analysis techniques. Based on this, it can be said that statistics has a connection with events in the past that may recur in the future. The statistical method that is often used to obtain future information is known as forecasting or predictive analysis. Predictive analysis is often employed in economics, finance, health, and other fields (Bakri et al., 2019; Ojugo & Yoro, 2020).

In the health sector, forecasting is often used as a means of evaluating the implementation, success and failure of a health program or health service that is being implemented. In addition, forecasting is also often used as a means of planning and decision making in the implementation of future activities. For example, Ranapurwala et al. (2019) used predictive modelling in the field of public health, namely agricultural vehicle accidents and concluded that forecasting or predictive data will be able to assist health policy makers (government, doctors, and health practitioners) in making decisions in an effort to improve public health; El Safty (2021) and El Safty & Al Zahrani (2021) used modelling in corona virus topic using topological method. On the other hand, forecasting methods can also be applied to several topics in the health sector, such as birth and death rates. The incidence of birth and death in an area is commonly used as an indicator in assessing the success of health services and health development programs in an area.

The infant mortality rate is one of the health problems in Indonesia that needs to be highlighted, because the infant birth rate is one of the indicators commonly used in determining public health. It is not surprising that health programs in Indonesia focus a lot on the problem of infant mortality, namely the reduction in infant mortality rates. In 2008, the Infant Mortality Rate in Indonesia was still quite high, around 31/1000, or in other words, 31 babies died in every 1,000 births. This mortality rate is higher when compared to Malaysia and Singapore, which amounted to 16.39/1000 and 2.3/1000 live births, respectively.

According to WHO data, in 2019, globally, as many as 7000 new-borns died every day and 185 cases per day occurred in Indonesia with an infant mortality rate of 24 per 1000 live births, with details of 75% of neonatal deaths occurring in the first week, and 40 % died within the first 24 hours (Jayanti et al., 2020). Given the importance of this infant mortality rate and to achieve one of the targets of the sustainable developmental goals (SDGs) in the health sector of the Republic of Indonesia, namely by 2030, to end preventable deaths of new-borns and children under five, with all countries trying to reduce the Neonatal Mortality Rate at least up to 12 per 1000 KH (Live Birth) and the under-five mortality rate of 25 per 1000, a suitable statistical method is needed in order to provide information in the future to minimize infant mortality. One of the statistical methods that is suitable for this problem of infant mortality is the method of prediction. Due to a decreasing trend from year to year, the infant mortality rate in Indonesia is assumed to meet the trend pattern. Forecasting methods that are suitable for the trend method are the ARIMA (Mishra et al., 2019), Neural Networks & Holt-Winters (Adeyinka & Muhajarine, 2020), and SutteARIMA (Ahmar & Boj, 2020a). SutteARIMA is used in this study because the SutteARIMA method is a new forecasting method that has a good level of accuracy in some forecasting data (Ahmar & Boj, 2020b).

3.2. Literature

3.2.1. ARIMA Method

The Autoregressive Integrate Moving Average (ARIMA) model was first discovered and presented by George Box and Gwilym Jenkins in 1976, and their names are often synonymous and associated with the ARIMA process applied for time series analysis, namely ARIMA Box-Jenkins. In general, the ARIMA model is written with the ARIMA notation (p, d, q), where p represents the order of the autoregressive process (AR), d represents the differencing, and q represents the order of the moving average (MA) process.

1) AR Process

The autoregressive model is a form of regression that connects the observed values at a certain time with the values of previous observations at certain intervals (Ahmar et al., 2013).

In general, the autoregressive process of data at the p level (AR (p)) (Wei, 2006):

$$\begin{gathered}Z_t = \varphi_1 Z_{t-1} + \varphi_2 Z_{t-2} + \ldots + \varphi_p Z_{t-p} + a_t, a_t \sim \mathrm{WN}\left(0, \sigma^2\right), \varphi_i \in \mathbf{R}, t \in \mathbf{Z}, \\ Z_t = \varphi_1 B Z_t + \varphi_2 B^2 Z_t + \ldots + \varphi_p B^p Z_t + a_t, a_t \sim \mathrm{WN}\left(0, \sigma^2\right), \varphi_i \in \mathbf{R}, t \in \mathbf{Z}, \\ (1 - \varphi_1 B + \varphi_2 B^2 + \ldots + \varphi_p B^p) Z_t = a_t, a_t \sim \mathrm{WN}\left(0, \sigma^2\right), \varphi_i \in \mathbf{R}, t \in \mathbf{Z}.\end{gathered} \tag{1}$$

This equation can be simplified into $\phi_{\mathbf{P}}(B) Z_t = a_t$, where: $\phi_{\mathbf{P}}(B) = 1 - \phi_1 B + \phi_2 B^2 + \ldots + \phi_p B^p$.

Because of this $\sum_{j=1}^{\infty}\left|\pi_j\right| = \sum_{j=1}^{p}\left|\phi_j\right| < \infty$ then this process is always invertible, for stationary, then the root of $\phi_{\mathbf{P}}(B) = 0$ must be outside the unit circle.

2) MA Process

The moving average process is a process that functions to describe phenomena in which the event produces an immediate effect which only lasts for a short period of time. The model of the general process moving average (MA) is as follows (Ahmar et al., 2013):

$$\begin{aligned}
Z_t &= a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \theta_q a_{t-1}, a_t \sim \mathrm{WN}\left(0,\sigma^2\right), \theta_i \in \mathbf{R}, t \in \mathbf{Z}, \\
Z_t &= a_t - \theta_1 B a_t - \theta_2 B^2 a_t - \ldots - \theta_q B^q a_t, a_t \sim \mathrm{WN}\left(0,\sigma^2\right), \theta_i \in \mathbf{R}, t \in \mathbf{Z}, \\
Z_t &= \left(1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q\right) a_t, a_t \sim \mathrm{WN}\left(0,\sigma^2\right), \theta_i \in \mathbf{R}, t \in \mathbf{Z}.
\end{aligned} \tag{2}$$

or

$$Z_t = \theta_Q(B) a_t, a_t \sim \mathrm{WN}\left(0,\sigma^2\right), \theta_i \in \mathbf{R}, t \in \mathbf{Z}. \tag{3}$$

with:

$$\theta_Q(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q. \tag{4}$$

Because $1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q < \infty$ then the finite moving average process is always stationary. This moving average process will be declared invertible if the root from $\theta_Q(B) = 0$ is outside of the circle.

3) ARMA Process

The model of the moving average autoregressive process (ARMA) (Ahmar et al., 2013):

$$\begin{aligned}
Z_t &= \varphi_1 Z_{t-1} + \varphi_2 Z_{t-2} + \ldots + \varphi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \theta_q a_{t-1}, \\
Z_t &= \varphi_1 B Z_t + \varphi_2 B^2 Z_t + \ldots + \varphi_p B^p Z_t + a_t - \theta_1 B a_t - \theta_2 B^2 a_t - \ldots - \theta_q B^q a_t, \\
(1 - \varphi_1 B + \varphi_2 B^2 + \ldots + \varphi_p B^p) Z_t &= \left(1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q\right) a_t, \\
\varphi_P(B) Z_t &= \theta_Q(B) a_t, a_t \sim \mathrm{WN}\left(0,\sigma^2\right), \varphi_i \in \mathbf{R}, t \in \mathbf{Z}.
\end{aligned} \tag{5}$$

with:

$$\varphi_{\mathbf{P}}(B) = 1 - \varphi_1 B + \varphi_2 B^2 + \ldots + \varphi_p B^p. \tag{6}$$

and

$$\theta_Q(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q. \tag{7}$$

For an invertible process, it is required that the root of $\theta_Q(B) = 0$ outside of the unit circle. And in order for it to be stationary, it is necessary that the root of $\phi_{\mathbf{P}}(B) = 0$ it outside of the unit circle. It is also assumed $\phi_{\mathbf{P}}(B) = 0$ and $\theta_Q(B) = 0$ don't have the same root. Furthermore, this process is referred to as the ARMA (p,q) model or process, where p and q are used to denote the respective order of polynomial values associated with autoregressive and moving averages.

4) ARIMA Process

The ARIMA process is basically similar to the ARMA process, they state that stationary and invertible processes can be represented in the form of a moving average or in an autoregressive form in the ARMA section. AR, MA, and ARMA require that data must be stationary, both in mean and in variance. Data can be stated as stationary in terms of average, if the time series data is relatively constant over time, it is stated to be stationary in variance, if the time series data structure from time to time has constant or constant data fluctuations and does not change or does not change the variance in the magnitude of the fluctuation. To overcome this non-stationary mean, a differencing process is carried out, and for non-stationary variants, a power transformation is carried out (λ). In the ARIMA modelling process, the variance stationarity process is carried out first then the average stationarity. From this stationary process comes the ARIMA process.

This ARIMA contains a differencing process to stationary data that is not stationary in the mean in the ARMA process. If there is a d-order differencing, then to achieve a stationary and general model of the ARIMA process (0, d, 0) it becomes:

$$(1-B)^d Z_t = a_t, a_t \sim \mathrm{WN}\left(0, \sigma^2\right), d, t \in \mathbf{Z}. \tag{8}$$

and from this equation, we can form a general model of the ARIMA process (p, d, q):

$$\varphi_P(B)(1-B)^d Z_t = \theta_Q(B) a_t, a_t \sim \mathrm{WN}\left(0, \sigma^2\right), \theta_i, \varphi_i \in \mathbf{R}, d, t \in \mathbf{Z}. \tag{9}$$

With the AR stationary operator $\phi_P(B)=1-\phi_1 B-\phi_2 B^2-\phi_p B^p$ and invertible MA operator $\theta_Q(B)=1-\theta_1 B-\theta_2 B^2-...-\theta_q B^q$ with no common factors.

5) The Stages in Forecasting the ARIMA(p, d, q)

In the execution of time Series data forecasting, the ARIMA method (p, d, q) have steps or stages. The stages in forecasting are as follows (Wei, 2006).

1) Model Identification

   Model identification is done to see the meaning of autocorrelation and data stationarity, to determine whether it is necessary to carry out a transformation or a differencing process (differentiation). From this stage, a temporary model will be obtained from which the process of testing the model will be carried out whether it is appropriate or not on the data.

2) Model Assessment and Testing

   After the model identification process has been carried out, the next step is to assess and test the model. This stage consists of two parts, namely parameter assessment and model diagnostic examination.

   a) Parameter Assessment

      After obtaining one or more provisional models, the next step is to find estimates for the parameters in that model.

   b) Model Diagnostic

      Diagnostic checking is done to check whether the estimated model is quite suitable or adequate with the existing data. Diagnostic checking is based on residual analysis. The basic assumption of the ARIMA model is that the residual is an independent random variable with a normal distribution with a constant mean of zero variance.

      (1) Independent Test

         This independent test is performed using the Box-Pierce Q statistical test. The Box-Pierce Q test can be calculated using the formula (Wei, 2006):

$$Q = n\sum_{k=1}^{m} \rho_k{}^2. \tag{10}$$

where:

n = amount of data

ρ_k = autocorrelation for lag k, k = 1,2, ..., m

If the value is $Q < \chi^2_{m-p-q}$, it is considered that the model is adequate, and vice versa, if the value is $Q > \chi^2_{m-p-q}$, it is considered inadequate.

(2) Normality Test

Residual analysis is used to examine whether the residuals of the model are white noise or not. White noise is the basic assumption of the ARIMA model where the residual in this case is a free random variable that is normally distributed with zero mean and constant variance.

3.2.2. Holt-Winters Method

Holt-Winters is a method for modeling and predicting the behavior of data from a time series. In addition, Holt-Winters is one of the most used time series forecasting methods. It is decades old but is still widely used in a variety of applications, including monitoring, which is used for things like anomaly detection and capacity planning. The Holt-Winters model uses three aspects of the time series: a typical value (average) / stationary, trend, and seasonality. Because it uses these three aspects, Holt-Winters is also known as triple exponential smoothing. Holt-Winters uses three smoothing parameters, namely α, β, γ, each of which has a value between 0 – 1.

The formula of Holt-Winters (Lestari et al., 2020):

Seasonal smoothing:

$$SN_t = \gamma(\frac{X_t}{S_t}) + (1-\gamma)SN_{t-L}. \tag{11}$$

Smooting data:

$$S_t = \alpha\frac{X_t}{SN_{t-1}} + (1-\alpha)(S_{t-1} + T_{t-1}). \tag{12}$$

Trend smooting:

$$T_t = \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1}. \tag{13}$$

Forecast:

$$F_{t+m} = (S_t + T_t m)SN_{t-L+m}, \tag{14}$$

$$F_{t+m} = (\alpha \frac{X_t}{SN_{t-1}} + (1-\alpha)(S_{t-1} + T_{t-1}) + \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1}m)\gamma(\frac{X_t}{S_t}),$$
$$F_{t+m} = +(1-\gamma)SN_{t-1}, \tag{15}$$

with:

X_t = actual data at time t,
S_t = data smoothing value,
S_{t-1} = previous data smoothing value,
T_t = trend smoothing
T_{t-1} = smoothing of the previous period's trend,
SN_t = seasonal smoothing,
SN_{t-1} = the previous period's seasonal smoothing,
a = exponential parameter for smoothing data with a value between 0 and 1,
β = exponential parameter for trend smoothing with a value between 0 and 1,
γ = exponential parameter for seasonal smoothing with a value between 0 and 1,
F_{t+m} = forecast value,
m = the time period to be predicted,
L = seasonal length.

3.2.3. Neural Network

An artificial neural network (ANN) is a system that processes information with characteristics and performance close to that of a biological neural network. Artificial neural networks are a generalization of biological neural network modeling with several assumptions, including:

a. Information processing lies in a number of components called neurons.
b. The signal spreads from one neuron to another via a connecting line.

c. Each connecting line has a weight and multiplies the value of the incoming signal (certain types of neurons).
d. Each neuron implements an activation function (usually nonlinear) which adds up all the inputs to determine the output signal.

Neural networks are useful for estimating or regression analysis including for forecasting and modelling, classification including pattern recognition and sequence recognition, as well as for decision making in sorting and processing data including filtering, grouping, and compression as well as programming of robots that move independently without human assistance. According to Wuryandari & Afrianto (2012), an artificial neural network model is determined by:

a. patterns of relationships between neurons (network architecture),
b. the method for determining and changing the joint weights is called the training method or network learning process,
c. activation function.

Artificial neural networks are also known as brain metaphors, computational neuron science, and parallel distributed processing. Neural networks are used for complex non-linear forecasting. One of the network requirements related to the time series is NNAR (Neural Network Autoregressive). Time series lag values can be used as input to neural networks, such as the lag values used in linear autoregressive models. This method is known as the neural network autoregressive model (NNAR). The NNAR model is generally denoted by NNAR(p, k) where p = input lag and k = number of hidden layers and NNAR (p, P, k) is the general denotation for NNAR in seasonality. For example, the NNAR(4,3) is a neural network that has four observational data ($y_{t-1}, y_{t-2}, \ldots, y_{t-4}$) which serve as input data used to predict the outcome or value of forecasting (Y_t) and is accompanied by three neurons in a hidden layer.

The NNAR model is a feed-forward neural network that involves a combination of linear and activation functions. This function formulation is defined as:

$$net_j = \sum_i w_{ij} y_{ij}, \tag{16}$$

$$f(y) = \frac{1}{1+e^{-y}}. \tag{17}$$

3.2.4. SutteARIMA Method

SutteARIMA is a short-term forecasting method developed by Ahmar and Boj in 2019 (Ahmar & Boj, 2020a). This method is a hybrid method between α-Sutte Indicator and ARIMA.

The formula of SutteARIMA (Ahmar & Boj, 2020a):

$$
\begin{aligned}
Z_t = {} & \alpha\left(\frac{\phi_1}{2}+\frac{\Delta z}{3\alpha+3\beta}\right)+\beta\left(\frac{\phi_3}{2}+\frac{2\Delta y}{3\beta+3\gamma}\right)+\gamma\left(\frac{\phi_3}{2}+\frac{2\Delta x}{3\gamma+3\delta}\right)+ \\
& \frac{\phi_4\delta}{2}+\ldots+\frac{\phi_p Z_{t-p}}{2}+\frac{a_t}{2}-\frac{\theta_1 a_{t-1}}{2}-\frac{\theta_2 a_{t-2}}{2}-\ldots-\frac{\theta_q a_{t-q}}{2}.
\end{aligned} \tag{18}
$$

3.3. Methodology and Data

3.3.1. Dataset

In this paper, we use annual time series data from Mortality rate, infant (per 1,000 live births) for Indonesia which is obtained from the World Bank Database. Data for this paper is available at: https://data.worldbank.org/indicator/SP.DYN.IMRT.IN?locations=ID. The World Bank website contains different annual time series at various levels of aggregation, from 1960 - 2019. In conducting data analysis, the data is divided into two parts, namely training data (from 1960-2012) and fitting / testing data (from 2013-2019). Training data is used to obtain forecasting models and fitting data is used to see the level of accuracy of the forecasting models obtained in the training data. Data were analyzed using forecasting methods: ARIMA, Neural Networks Time Series, Holt-Winters, and SutteARIMA method. To simplify the analysis, we used the R Software version 3.6.3, namely the SutteForecastR package (Ahmar, 2017) and Microsoft Excel 2010.

3.3.2. Forecast Accuracy

In the results of the fitting / testing data, two performance indicators or forecasting accuracy are used to assess the quality with the good of fit standard and the accuracy of the forecasting results obtained. The indicators are as follows (Ahmar, 2020).

- Mean absolute percentage error, MAPE

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \tag{19}$$

- Mean square error, MSE

$$\text{MSE} = \frac{\sum_{t=1}^{n}\left(A_t - F_t\right)^2}{n} \tag{20}$$

where:

A_t = Actual values at data time t.
F_t = Forecast values at data time t.

3.4. Results and Discussion

In the case of infant mortality rates, the data obtained is in the form of a trend and has decreased every year (see figure 3.2). Figure 3.2. shows that the data on infant mortality in 2003 and 2004 were constant and then in 2005 decreased slowly until 2019.

3.4.1. Model Specification

To obtain forecasting models and forecasting results from data using ARIMA, Neural Network Time Series, Holt-Winter, and SutteARIMA models, we use the *alpha.sutte* function on SutteForecastR package on Software and the output results from R Software are presented as follows.

1) Forecasting Using ARIMA Method

$Forecast_AutoARIMA

Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
23	25.05055	24.44849	25.65261	24.12978	25.97132
24	24.11898	23.31295	24.92502	22.88626	25.35171
25	23.18742	22.07214	24.30271	21.48174	24.89310
26	22.25586	20.75409	23.75763	19.95910	24.55262
27	21.32430	19.37637	23.27223	18.34520	24.30340
28	20.39274	17.94907	22.83641	16.65547	24.13001
29	19.46118	16.47840	22.44395	14.89941	24.02294

$AutoARIMA
Series: al_mi_10

ARIMA(0,2,2)

Coefficients:

	ma1	ma2
	-1.1098	0.5000
s.e.	0.2471	0.2149

sigma^2 estimated as 0.2207: log likelihood=-12.88
AIC=31.76 AICc=33.26 BIC=34.75

From the results of the analysis output, it is obtained the ARIMA model (0,2,2) with 2 times differencing and the values of MA (1): -1.1098 and MA (2): 0.5000. The form of the model is as follows.

$$Z_t = 1.1098a_{t-1} - 0.0.5000a_{t-2} + \varepsilon.$$

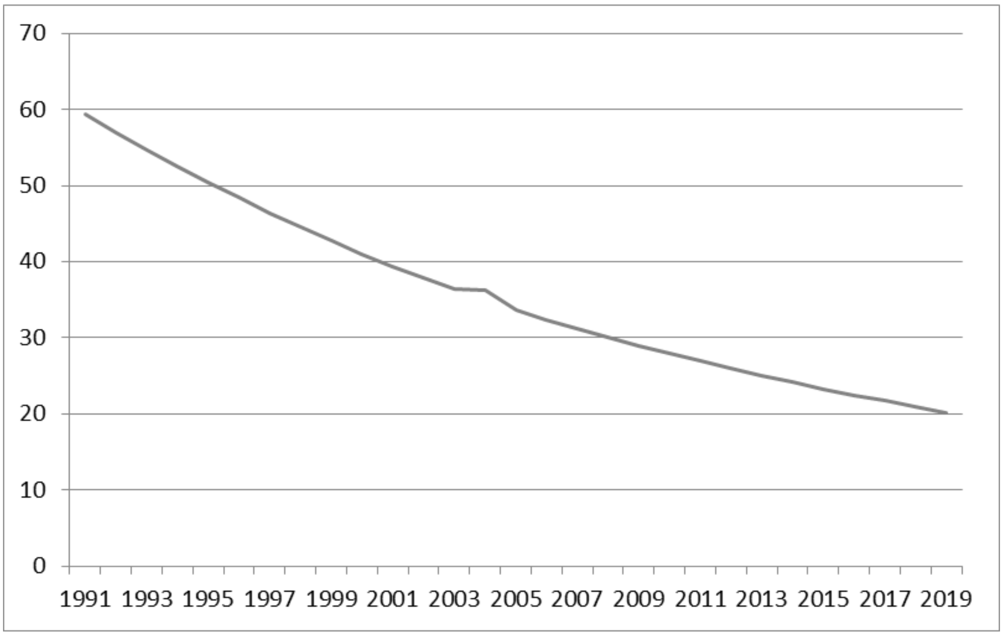


Figure 3.2. Infant Mortality Rates in Indonesia

2) Forecasting Using Neural Network Time Series Method

$Forecast_NNETAR

Point	Forecast
23	25.40859
24	24.90951
25	24.49229

26	24.14630
27	23.86136
28	23.62803
29	23.43788

$NNETAR
Series: al_mi_10
Model: NNAR(1,1)
Call: nnetar(y = al_mi_10)

Average of 20 networks, each of which is
a 1-1-1 network with 4 weights
options were - linear output units

sigma^2 estimated as 0.1567

Like the ARIMA method, the Neural Network Time Series method is obtained by a forecasting model, namely NNAR(1,1) with 1 hidden screen.

3) Forecasting Using Holt-Winters Method

$Forecast_HoltWinters

Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
23	25.05943	24.53140	25.58745	24.25188	25.86698
24	24.15237	23.47042	24.83433	23.10941	25.19533
25	23.24532	22.31579	24.17485	21.82372	24.66691
26	22.33826	21.09096	23.58556	20.43068	24.24584
27	21.43120	19.81246	23.04994	18.95555	23.90685
28	20.52415	18.49000	22.55830	17.41319	23.63511
29	19.61709	17.12949	22.10469	15.81264	23.42154

$HoltWinters
Holt-Winters exponential smoothing with trend and without seasonal component.

Call:
HoltWinters(x = al_mi_10, gamma = FALSE)

Smoothing parameters:
 alpha: 0.4384154
 beta : 0.8642498

gamma: FALSE
Coefficients:
[,1]
a 25.9664829
b -0.9070558

From the analysis for forecasting using the Holt-Winters method (table 3.3), the Holt-Winters method for this data on infant mortality is Holt-Winters without seasonal components or using only two parameters, namely $\alpha = 0.4384154$ and $\beta = 0.8642498$.

4) Forecasting Using SutteARIMA Method

$Tes_Data
[1] 25.1 24.2 23.3 22.5 21.7 20.9 20.2

$Forecast_AlphaSutte
[1] 25.01820 24.15097 23.28372 22.41644 21.64915 20.88184 20.11449
$Forecast_AutoARIMA

Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
23	25.05055	24.44849	25.65261	24.12978	25.97132
24	24.11898	23.31295	24.92502	22.88626	25.35171
25	23.18742	22.07214	24.30271	21.48174	24.89310
26	22.25586	20.75409	23.75763	19.95910	24.55262
27	21.32430	19.37637	23.27223	18.34520	24.30340
28	20.39274	17.94907	22.83641	16.65547	24.13001
29	19.46118	16.47840	22.44395	14.89941	24.02294

$Forecast_SutteARIMA

Point	Forecast	Low 95	High 95
23	25.0344	23.7827	26.2861
24	24.1350	22.9282	25.3417
25	23.2356	22.0738	24.3973
26	22.3362	21.2193	23.4530
27	21.4867	20.4124	22.5611
28	20.6373	19.6054	21.6692
29	19.7878	18.7984	20.7772

3.4.2. Estimating the Forecasting Model

After obtaining the forecasting model in the specification model section, the result of forecasting for testing data are shown in table 3.1 to table 3.4. The forecasting results are compared with the testing data to obtain the value of absolute percentage error (APE) and square of error (SE).

Table 3.1. Forecasting results for data testing with the ARIMA(0,2,2)

Data	Forecast	Low 95	High 95	APE	SE
25.1	25.05055	24.12978	25.97132	0.001293059	0.004307
24.2	24.11898	22.88626	25.35171	0.001324584	0.004228
23.3	23.18742	21.48174	24.89310	0.004135937	0.004151
22.5	22.25586	19.95910	24.55262	0.007163493	0.026847
21.7	21.32430	18.34520	24.30340	0.015005208	0.045486
20.9	20.39274	16.65547	24.13001	0.023422265	0.069017
20.2	19.46118	14.89941	24.02294	0.032479571	0.169880
			Mean	**1.211773099**	**0.046274**

Table 3.2. Forecasting results for data testing with NNAR(1,1)

Data	Forecast	Low 95	High 95	APE	SE
25.1	25.40859	24.13816	26.67902	0.012294422	0.095228
24.2	24.90951	23.66403	26.15499	0.029318595	0.503404
23.3	24.49229	23.26768	25.7169	0.051171245	1.421555
22.5	24.14630	22.93899	25.35362	0.073168889	2.710304
21.7	23.86136	22.66829	25.05443	0.099601843	4.671477
20.9	23.62803	22.44663	24.80943	0.130527751	7.442148
20.2	23.43788	22.26599	24.60977	0.160291089	10.483870
			Mean	**7.948198**	**3.903998**

Table 3.3 Forecasting results for data testing with Holt-Winters

Data	Forecast	Low 95	High 95	APE	SE
25.1	25.05943	24.25188	25.86698	0.001616335	0.001646
24.2	24.15237	23.10941	25.19533	0.001968182	0.002269
23.3	23.24532	21.82372	24.66691	0.002346781	0.002990
22.5	22.33826	20.43068	24.24584	0.007188444	0.026160
21.7	21.43120	18.95555	23.90685	0.012387097	0.072253

Data	Forecast	Low 95	High 95	APE	SE
20.9	20.52415	17.41319	23.63511	0.017983254	0.141263
20.2	19.61709	15.81264	23.42154	0.028856931	0.339784
			Mean	**1.033528901**	**0.083766**

Table 3.4. Forecasting results for data testing with SutteARIMA

Data	Forecast	Low 95	High 95	APE	SE
25.1	25.0344	23.7827	26.2861	0.002615	0.004307
24.2	24.1350	22.9282	25.3417	0.002687	0.004228
23.3	23.2356	22.0738	24.3973	0.002765	0.004151
22.5	22.3362	21.2193	23.4530	0.007282	0.026847
21.7	21.4867	20.4124	22.5611	0.009828	0.045486
20.9	20.6373	19.6054	21.6692	0.012570	0.069017
20.2	19.7878	18.7984	20.7772	0.020404	0.169880
			Mean	**0.830734**	**0.046274**

3.4.3. Model's Forecasting Performance Comparison

Based on the results of forecasting for testing data from various prediction methods that have been presented previously, the comparison of the forecasting results is presented in figure 3.3. In figure 3.3, it can be seen that SutteARIMA has the highest level of accuracy in predicting infant mortality rates. This was followed by the Holt-Winters, ARIMA(0,2,2), and NNAR(1,1). The results of the forecasting graphs of each forecasting method for testing data are presented in Figure 3.4. From Figure 3.4., it can be seen that the ARIMA(0,2,2), Holt-Winters, and SutteARIMA methods go hand in hand as shown in Figure 4 which shows MSE and MAPE results. In fact, these three methods are close to or not too far away and differ from the NNAR(1,1) whose forecasting is inaccurate with the testing data. Based on these results, SutteARIMA is used as a method for forecasting the next 5 periods or years (table 3.5).

Table 3.5. Forecasting the 5 Future Periods

Year	Forecast
2020	19.7557
2021	19.0523
2022	18.7566

Year	Forecast
2023	18.0748
2024	17.9185

Based on the forecast results in table 3.5, there is a decrease in the infant mortality rate from year to year. This result is in line with the opinion of Mishra et al. (2019); Kurniasih et al. (2018), and Hussein (1991) who said that the infant mortality rate has decreased from year to year.

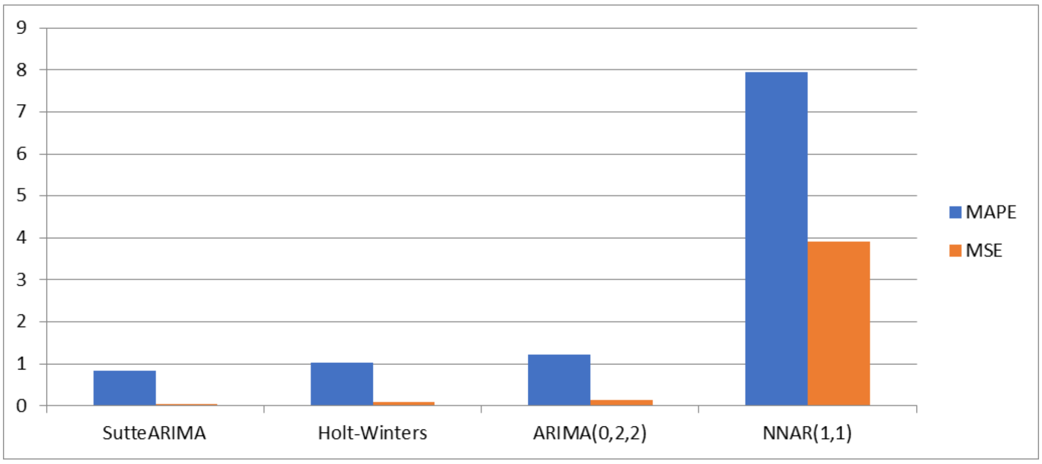


Figure 3.3. Comparison of Forecasting Accuracy on Infant Mortality Rate in Indonesia.

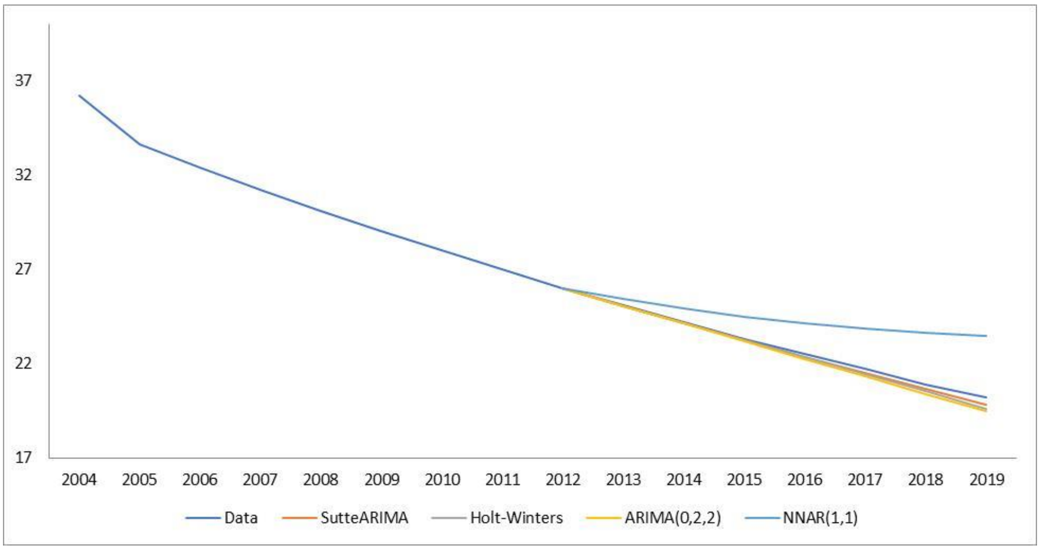


Figure 3.4. Testing Data Forecasting Results on Infant Mortality Rates in Indonesia.

3.5. Conclusion and Impact

The purpose of this research is to model the infant mortality rate data and find the best model to predict this problem in the future. To achieve this goal, four model are used (ARIMA, Holt-Winters, Neural Network Time Series, and SutteARIMA) to predict the infant mortality rate data. To determine which prediction model is more suitable and precise in predicting data, the MAPE and MSE values of each of the forecasting methods used are calculated and the results are compared according to the predetermined performance criteria. Based on the findings of this study, it is concluded that the better or more suitable model, with smaller forecast errors in the infant mortality case data, is SutteARIMA which is then followed by Holt-Winters, ARIMA, and NNAR. And based on data trends and forecast results, the infant mortality rate is decreasing from year to year. The SutteARIMA method provides an estimated infant mortality rate for 2020 of 19.7557 and 17.9185 for 2024, a decline from 2019. These findings have the potential to help promote policies in order to address and minimize infant mortality rates in the coming years and can be used as a basis for implementing appropriate strategies to overcome them so that Indonesia's SDGs targets can be achieved. Although the infant mortality rate is predictable and has a satisfactory level of accuracy, it is possible that the results of this prediction are not precise due to human behaviour and policies taken by policy makers.

Chapter 4
SutteARIMA: A Novel Forecasting Method for Financial and Economic Data and its Comparison with Other Forecasting Methods[1,2]

[1] This chapter presented at International Conference on Social, Economics, Business, and Education 2021, Indonesia; 2nd International Conference on Sciences and Mathematics, 2021, Indonesia

[2] This chapter under review in Journal of The Franklin Institute.

4.1. Introduction

Indonesia and all countries in the world are currently hit by a big problem, namely the problem of the COVID-19 pandemic. COVID-19 is a pandemic disease caused by the SAR-Cov virus. The COVID-19 pandemic has had a huge impact and disrupted all aspects of human life. Not only the educational aspect but also the economic aspect. The inhibition of economic activity automatically makes business actors perform efficiency to reduce losses. As a result, many workers are furloughed or even laid off (PHK). Based on data from the Ministry of Manpower of Republic of Indonesia as of April 7, 2020, due to the COVID-19 pandemic, there were 39,977 companies in the formal sector that chose to lay off and lay off their workers. A total of 1,010,579 workers were affected by this (Rizal, 2020).

Accurate forecasting can be used as a strong basis for making better plans regarding solutions to economic problems, especially during the COVID-19 pandemic (Ahmar and Boj, 2020). Of course, an efficient and accurate forecasting technique is needed. The time series forecasting method is one of the important forecasts because past observations are collected and analyzed and used in developing forecasting models that can explore future time series (Kumar and Vanajakshi, 2015). In the last few decades, many forecasting methods have been developed.

In the data forecasting method, time series data forecasting can be divided into two parts, namely linear methods and non-linear methods. Linear methods that are often used in forecasting are the nave method (Goodwin, 2014; Qeadan et al., 2020), the exponential smoothing model (Billah et al., 2006), the Holt-Winters model (Tratar and Ahmar, 2019), and the ARIMA model (Abonazel and Abd-Elftah, 2019; Ranjbar and Khaledian, 2014). Among these linear models, the ARIMA model is a reliable forecasting model and has been successfully tested in various fields. In its use, the user must determine a suitable model without the need for knowledge of the complexity of the data. Unfortunately, the model of this linear method is suitable only for linear data and of course there will be a mismatch of forecasting results if it turns out that the data used is non-linear data so that non-linear models need to be considered.

In non-linear forecasting models, the methods often used by researchers are the Neural Network (NN) and support vector regression (SVR) methods (Moghaddam et al., 2016; Yao et al., 2015). The NN and SVR methods are designed to study nonlinear patterns of a time series. This method is generally believed to outperform linear methods in economic modeling and others and is more efficient in assisting decision making.

In practice, there are many stationary phenomena that can be described by linear time series methods and many non-linear data phenomena that can be described by non-linear time series methods. But this phenomenon does not always occur in forecasting. For example, (Pemberton, 1989) and (Gokcan, 2000) found that non-linear forecasting methods do not always have better forecasting performance than linear methods. So for this problem, Ahmar et al. (2018) developed a forecasting method, namely the -Sutte Indicator which is claimed to be a forecasting method that can be used regardless of whether the data is linear or linear in the sense that the assumptions needed so far in carrying out a Forecasting is not required because this method uses forecasting methods using previous data. Many researchers compare the -Sutte Indicator method with linear and non-linear methods in their research and reveal that the -Sutte Indicator method is more efficient than other models. Furthermore, in -Sutte the indicator is only suitable for data that has the characteristics of Up Trend & Down Trend as well as Non Seasonal/Cyclic (Ahmar, 2019).

A good forecasting method is a forecasting method that has a good level of forecasting accuracy or in the sense that the value of the difference between the forecasting results and the original data is small. Forecasting methods have been developed by various researchers, but these methods are limited in their accuracy and lack of precision with the original data. So based on this phenomenon, the author then develops a forecasting method that combines the phenomena that were previously explained. This method is known as SutteARIMA. The combination of forecasting methods has been suggested by various researchers, for example: Aiolfi and Timmermann, (2006) used the new four-stage conditional model combination method and found the results that the proposed method worked well on the forecasting trial sample, furthermore, Pauwels and Vasnev (2014) also examined the combination of forecasting on the US business cycle index and get the result that the combination method increases the level of forecasting accuracy.

SutteARIMA has conducted simulation trials on 200 datasets generated in the form of non-trend, trend, non-seasonal, and seasonal data and the results are suitable for (1) Down Trend, Non-Seasonal/Cyclic data; (2) Up & Down Trend, Seasonal/Cyclic; (3) Trend, Seasonal/Cyclic; and (4) Up & Down Trend, Seasonal/Cyclic; or in other words SutteARIMA can be used on data containing linear, non-linear data or a combination of linear and non-linear (Ahmar, 2019), more details will be presented in the SutteARIMA approach section.

In this paper, we propose a combined forecasting approach that combines linear and non-linear data. The approach method in question is named SutteARIMA. The SutteARIMA method will be tested in the case of National Currency (IDR) to US Dollar Spot Exchange Rate for Indonesia and Consumer Price Index: All Items for Indonesia. This is because these two cases use linear and non-linear data.

To see the level of accuracy of SutteARIMA and other forecasting methods (linear and non-linear), then in the results and discussion section a comparison of forecasting results between SutteARIMA and other forecasting methods using MAPE and MSE will be explained which will be explained in the methodology section. It is hoped that with this comparison of accuracy levels, knowledge and related input can be obtained regarding the proposed combination forecasting method.

4.2. SutteARIMA approach

SutteARIMA is a forecasting method that combines the -Sutte Indicator and ARIMA methods. The following describes the derivation of the formula from SutteARIMA (Ahmar & Boj, 2020).

The formula for the α-Sutte Indicator (Ahmar, 2018):

$$Z_t = \frac{\gamma\left(\dfrac{\Delta x}{\dfrac{\gamma+\delta}{2}}\right)+\beta\left(\dfrac{\Delta y}{\dfrac{\beta+\gamma}{2}}\right)+\alpha\left(\dfrac{\Delta z}{\dfrac{\alpha+\beta}{2}}\right)}{3}, \tag{4.1}$$

where:

$\delta = Z_{t-4}$,
$\gamma = Z_{t-3}$,
$\beta = Z_{t-2}$,
$\alpha = Z_{t-1}$,
$\Delta x = \gamma - \delta = Z_{t-3} - Z_{t-4}$,
$\Delta y = \beta - \gamma = Z_{t-2} - Z_{t-3}$,
$\Delta z = \alpha - \beta = Z_{t-1} - Z_{t-2}$,
Z_t = data at t time,
Z_{t-k} = data at (t – k) time.

Equation (4.1) can be simplified to:

$$Z_t = \frac{\gamma\left(\dfrac{\Delta x}{\dfrac{\gamma+\delta}{2}}\right)+\beta\left(\dfrac{\Delta y}{\dfrac{\beta+\gamma}{2}}\right)+\alpha\left(\dfrac{\Delta z}{\dfrac{\alpha+\beta}{2}}\right)}{3},$$

$$Z_t = \frac{\dfrac{\gamma\Delta x}{\dfrac{\gamma+\delta}{2}}+\dfrac{\beta\Delta y}{\dfrac{\beta+\gamma}{2}}+\dfrac{\alpha\Delta z}{\dfrac{\alpha+\beta}{2}}}{3},$$

$$Z_t = \frac{\gamma\Delta x}{\frac{3\gamma+3\delta}{2}} + \frac{\beta\Delta y}{\frac{3\beta+3\gamma}{2}} + \frac{\alpha\Delta z}{\frac{3\alpha+3\beta}{2}},$$

$$Z_t = \frac{2\gamma\Delta x}{3\gamma+3\delta} + \frac{2\beta\Delta y}{3\beta+3\gamma} + \frac{2\alpha\Delta z}{3\alpha+3\beta},$$

$$Z_t = \gamma\frac{2\Delta x}{3\gamma+3\delta} + \beta\frac{2\Delta y}{3\beta+3\gamma} + \alpha\frac{2\Delta z}{3\alpha+3\beta}, \tag{4.2}$$

and the formula for ARIMA(p,d,q):

$$\phi_p(B)Z_t = \theta_q(B)a_t, a_t \sim WN\left(0,\sigma^2\right), \phi_p, \theta_q \in \mathbf{R}, \mathrm{t} \in \mathbf{Z}, \tag{4.3}$$

with $\phi_p(B) = (1-\phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p)$ (for AR(p))

and $\theta_q(B) = (1-\theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q)$ (for MA(q)).

Equation (4.3.) can be further elaborated and obtained:

$$\begin{aligned} &(1-\phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p)Z_t = (1-\theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q)a_t, \\ &Z_t - \phi_1 BZ_t - \phi_2 B^2 Z_t - \ldots - \phi_p B^p Z_t = a_t - \theta_1 Ba_t - \theta_2 B^2 a_t - \ldots - \theta_q B^q a_t. \end{aligned} \tag{4.4}$$

The backward shift operator equation $B^p Z_t$ can be changed to Z_{t-p}. If equation (4.4) is changed according to the backward shift operator equation, then we get:

$$\begin{aligned} &Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2} - \ldots - \phi_p Z_{t-p} = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \theta_q a_{t-q}, \\ &Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots + \phi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \theta_q a_{t-q}. \end{aligned} \tag{4.5}$$

If we define:

$$\begin{aligned} &\delta = Z_{t-4}, \\ &\gamma = Z_{t-3}, \\ &\beta = Z_{t-2}, \\ &\alpha = Z_{t-1}. \end{aligned} \tag{4.6}$$

We substitute equation (4.6) to the equation (4.5):

$$Z_t = \phi_1\alpha + \phi_2\beta + \phi_3\gamma + \phi_4\delta + \ldots + \phi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \theta_q a_{t-q}. \tag{4.7}$$

In the last stage (combination of formulas), we do the addition process between equations (4.2) and (4.7), so that we get:

$$
\begin{aligned}
2Z_t &= \phi_1\alpha + \phi_2\beta + \phi_3\gamma + \phi_4\delta + \ldots + \phi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \theta_q a_{t-q} + \\
&\quad \gamma\frac{2\Delta x}{3\gamma + 3\delta} + \beta\frac{2\Delta y}{3\beta + 3\gamma} + \alpha\frac{2\Delta z}{3\alpha + 3\beta}, \\
Z_t &= \alpha\left(\frac{\phi_1}{2} + \frac{\Delta z}{3\alpha + 3\beta}\right) + \beta\left(\frac{\phi_3}{2} + \frac{2\Delta y}{3\beta + 3\gamma}\right) + \gamma\left(\frac{\phi_3}{2} + \frac{2\Delta x}{3\gamma + 3\delta}\right) + \\
&\quad \frac{\phi_4\delta}{2} + \ldots + \frac{\phi_p Z_{t-p}}{2} + \frac{a_t}{2} - \frac{\theta_1 a_{t-1}}{2} - \frac{\theta_2 a_{t-2}}{2} - \ldots - \frac{\theta_q a_{t-q}}{2}.
\end{aligned} \tag{4.8}
$$

Equation (4.8) is the final formula of SutteARIMA.

The SutteARIMA forecasting method will conduct simulation trials on 200 datasets generated in the form of non-trend, trend, non-seasonal, and seasonal data, and will be compared with various forecasting methods that already exist including -Sutte, ARIMA, Holt-Winter, NNETAR, Robust, and Theta. The results of the forecasting are obtained from the output of R Software using forecast (Hyndman, et.al., 2021), fracdiff (Maechler, 2020), robets (Crevits, Bergmeir, & Hyndman, 2018), forecastHybrid (Shaub & Ellis, 2020) packages. The results of this trial are described as follows.

1) The 1st Generated Artificial Data (Non Trend, Non Seasonal/Cyclic)

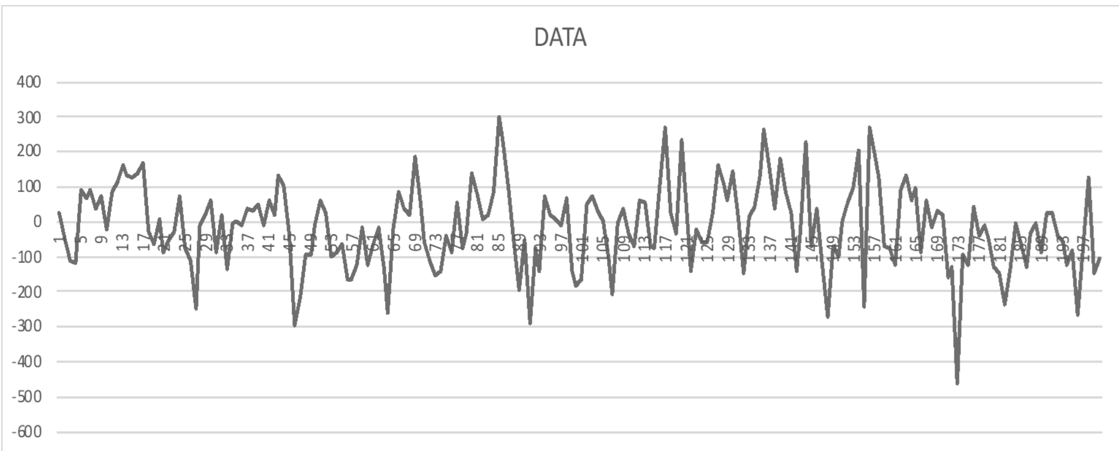


Figure 4.1. Plot 1st Generated Artificial Data (Non Trend, Non Seasonal/Cyclic)

Table 4.1. Forecasting accuracy of 1st Generated Artificial Data (Non Trend, Non Seasonal/Cyclic)

Data	Forecast						
	α-Sutte	ARIMA	Holt-Winter	NNETAR	Robust	Theta	Sutte ARIMA
-121.31	-156.11	-24.23	-41.55	-51.13	-38.61	-37.99	-90.17
-81.18	-248.13	-10.20	-41.32	-48.08	-38.48	-38.04	-129.16
-268.56	-105.26	-4.30	-41.09	-46.54	-38.35	-38.10	-54.78
-67.64	-382.59	-1.81	-40.86	-45.72	-38.22	-38.16	-192.20
128.97	-125.89	-0.76	-40.63	-45.27	-38.10	-38.21	-63.33
-148.37	335.63	-0.32	-40.40	-45.00	-37.98	-38.27	167.66
-106.96	-1259.70	-0.13	-40.17	-44.85	-37.86	-38.33	-629.92
Forecast accuracy							
MSE	254712.42	19827.23	15041.67	**14384.68**	15342.36	15366.60	67833.16
MAPE	337.48	94.78	72.27	**68.08**	74.08	74.18	171.37
MAE*	367.37	126.11	102.60	**98.13**	104.51	104.62	206.96

* Mean absolute error (MAE)

2) The 2nd Generated Artificial Data (Non Trend, Non Seasonal/Cyclic)

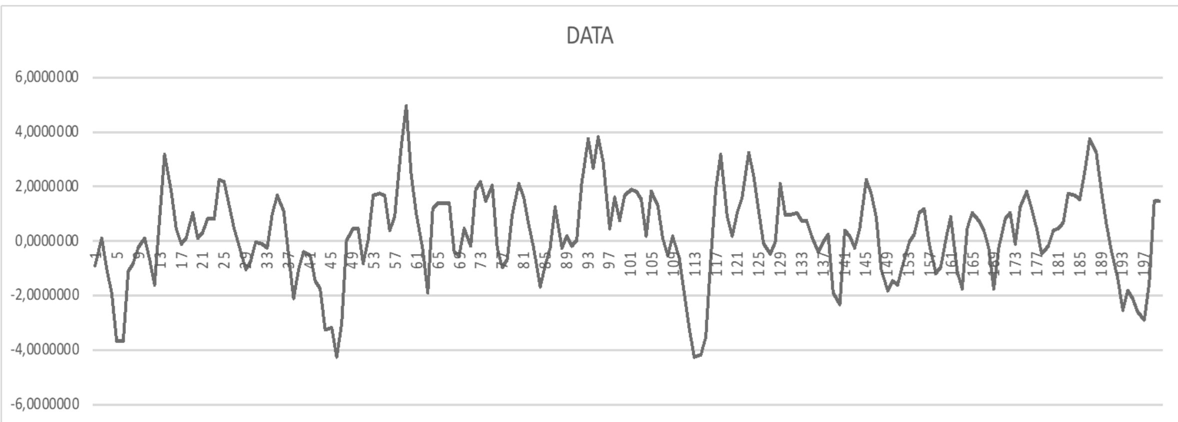


Figure 4.2. Plot 2nd Generated Artificial Data (Non Trend, Non Seasonal/Cyclic)

Table 4.2. Forecasting accuracy of 2nd Generated Artificial Data (Non Trend, Non Seasonal/Cyclic)

Data	Forecast						
	α-Sutte	ARIMA	Holt-Winter	NNETAR	Robust	Theta	Sutte ARIMA
0.11	-0.45	-0.16	-0.31	0.25	-0.37	0.19	-0.31
0.49	0.46	0.00	-0.36	0.24	-0.37	0.19	0.23
-1.37	0.87	0.00	-0.41	0.20	-0.37	0.19	0.44
0.52	-2.92	0.00	-0.46	0.18	-0.37	0.19	-1.46
2.52	-1.95	0.00	-0.51	0.18	-0.37	0.19	-0.98
-0.05	0.95	0.00	-0.56	0.18	-0.37	0.19	0.47
-1.26	0.33	0.00	-0.61	0.19	-0.37	0.19	0.16
Forecast accuracy							
MSE	5.82	1.49	1.80	1.47	1.71	**1.46**	3.14
MAPE	507.36	122.13	277.73	145.50	234.90	**140.78**	314.74
MAE	1.91	0.93	1.05	0.90	1.04	**0.90**	1.42

3) The 3rd Generated Artificial Data (Non Trend, Non Seasonal/Cyclic)

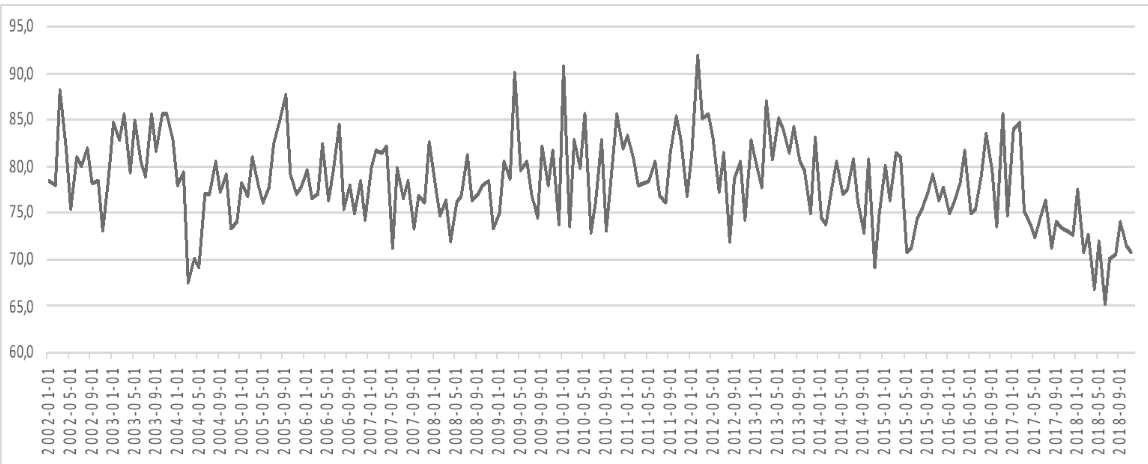


Figure 4.3. Plot 3rd Generated Artificial Data (Non Trend, Non Seasonal/Cyclic)

Table 4.3. Forecasting accuracy of 3rd Generated Artificial Data (Non Trend, Non Seasonal/Cyclic)

Data	Forecast						
	α-Sutte	ARIMA	Holt-Winter	NNETAR	Robust	Theta	Sutte ARIMA
71.90	63.29	72.64	71.50	72.47	72.30	72.75	67.96
65.10	72.42	72.64	71.36	72.66	72.30	72.74	72.53
70.10	62.89	72.64	71.21	73.45	72.30	72.73	67.77
70.40	71.47	72.64	71.07	74.67	72.30	72.73	72.06
74.00	70.07	72.64	70.93	77.74	72.30	72.72	71.36
71.40	77.06	72.64	70.78	78.41	72.30	72.71	74.85
70.80	71.88	72.64	70.64	78.44	72.30	72.70	72.26
Forecast accuracy							
MSE	32.78	10.81	**7.27**	29.79	9.49	11.21	14.27
MAPE	7.11	3.66	**2.57**	6.99	3.31	3.75	4.72
MAE	4.98	2.50	**1.76**	4.88	2.26	2.56	3.27

From the test results on non-trend and non-seasonal/cyclic data (table 4.1, table 4.2, and table 4.3), it was found that the NNETAR, Theta, and Holt-Winter methods have a good level of accuracy based on the MSE, MAPE, and MAE values.

4) The 1st Generated Artificial Data (Up Trend, Non Seasonal/Cyclic)

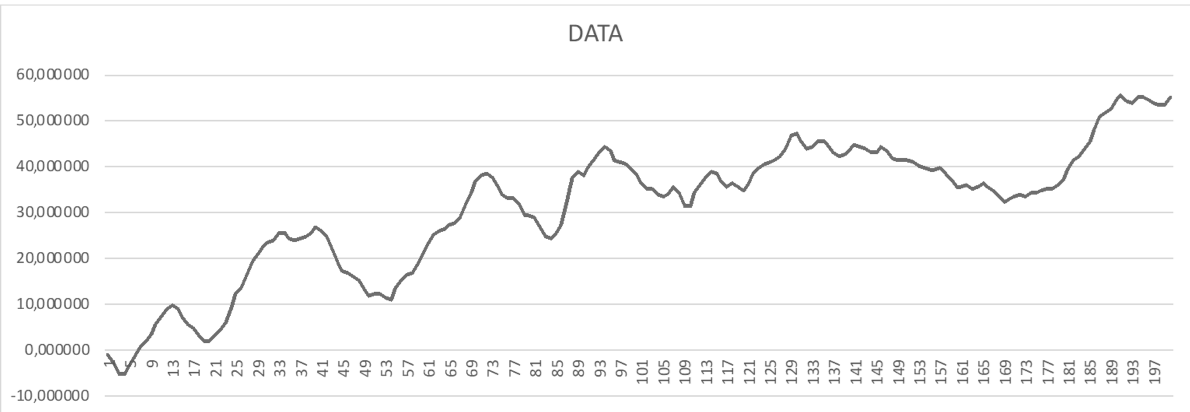


Figure 4.4. Plot 1st Generated Artificial Data (Up Trend, Non Seasonal/Cyclic)

Table 4.4. Forecasting accuracy of 1st Generated Artificial Data (Up Trend, Non Seasonal/Cyclic)

Data	Forecast						
	α-Sutte	ARIMA	Holt-Winter	NNETAR	Robust	Theta	Sutte ARIMA
54.84	53.62	54.13	53.50	54.14	53.75	53.99	53.87
55.20	54.65	53.94	53.11	54.58	53.64	54.10	54.30
54.61	55.51	53.26	52.73	54.92	53.54	54.20	54.38
53.65	54.86	52.56	52.34	55.10	53.45	54.31	53.71
53.23	53.26	52.03	51.95	55.19	53.37	54.41	52.64
53.29	52.57	51.51	51.57	55.27	53.31	54.52	52.04
55.12	52.86	50.95	51.18	55.37	53.25	54.63	51.90
Forecast accuracy							
MSE	1.39	3.87	4.51	1.56	1.19	**0.82**	2.01
MAPE	1.81	3.04	3.56	1.93	**1.55**	1.57	1.89
MAE	0.99	1.65	1.94	1.04	**0.85**	**0.85**	1.03

5) The 2nd Generated Artificial Data (Up Trend, Non Seasonal/Cyclic)

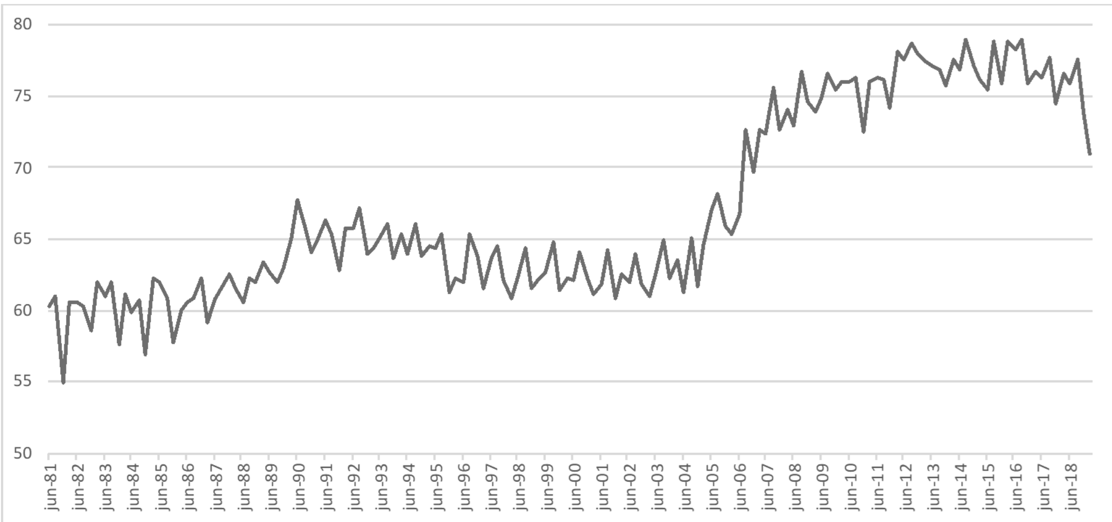


Figure 4.5. Plot 2nd Generated Artificial Data (Up Trend, Non Seasonal/Cyclic)

Table 4.5. Forecasting accuracy of 2nd Generated Artificial Data (Up Trend, Non Seasonal/Cyclic)

Data	Forecast						
	α-Sutte	ARIMA	Holt-Winter	NNETAR	Robust	Theta	Sutte ARIMA
77.70	75.42	77.13	76.98	77.43	77.15	76.92	76.28
74.50	78.34	77.25	76.87	75.89	77.09	76.98	77.80
76.50	73.79	77.37	76.75	76.83	77.03	77.05	75.58
75.90	76.60	77.49	76.64	76.44	76.99	77.11	77.05
77.50	75.33	77.62	76.53	77.12	76.95	77.18	76.47
73.80	78.52	77.74	76.41	76.23	76.92	77.24	78.13
70.90	72.94	77.86	76.30	76.90	76.90	77.31	75.40
Forecast accuracy							
MSE	8.40	10.73	**6.24**	6.34	7.78	8.80	7.86
MAPE	3.51	3.28	2.55	**2.23**	2.81	2.97	3.21
MAE	2.63	2.40	1.87	**1.62**	2.06	2.17	2.38

From the test results on up-trend and non-seasonal/cyclic data (table 4.4 and table 4.5), it was found that the Theta and Holt-Winter methods have a good level of accuracy based on MSE values and Robust and NNETAR for MAPE and MAE values.

6) The Generated Artificial Data (Down Trend, Non Seasonal/Cyclic)

Figure 4.6. Plot Generated Artificial Data (Down Trend, Non Seasonal/Cyclic)

Table 4.6. Forecasting accuracy of Generated Artificial Data (Down Trend, Non Seasonal/Cyclic)

Data	Forecast						
	α-Sutte	ARIMA	Holt-Winter	NNETAR	Robust	Theta	Sutte ARIMA
-74.17	-73.99	-73.83	-74.43	-73.66	-74.00	-73.57	-73.91
-75.45	-74.88	-74.60	-75.60	-73.90	-74.54	-73.75	-74.74
-75.22	-76.58	-74.99	-76.77	-74.11	-74.97	-73.94	-75.79
-75.63	-75.83	-75.38	-77.94	-74.29	-75.31	-74.12	-75.60
-76.36	-76.13	-75.76	-79.11	-74.44	-75.58	-74.30	-75.94
-76.24	-76.66	-76.15	-80.28	-74.58	-75.80	-74.49	-76.40
-78.08	-76.59	-76.54	-81.46	-74.69	-75.98	-74.67	-76.56
Forecast accuracy							
MSE	0.68	0.53	6.16	3.38	0.90	3.74	**0.49**
MAPE	0.83	0.73	2.70	2.15	0.93	2.30	**0.69**
MAE	0.64	0.56	2.06	1.64	0.71	1.76	**0.52**

From the test results on down trend and non-seasonal/cyclic data (table 4.6), it is found that the SutteARIMA method has a good level of accuracy based on the MSE, MAPE, and MAE values.

7) The Generated Artificial Data (Up & Down Trend, Non Seasonal/Cyclic)

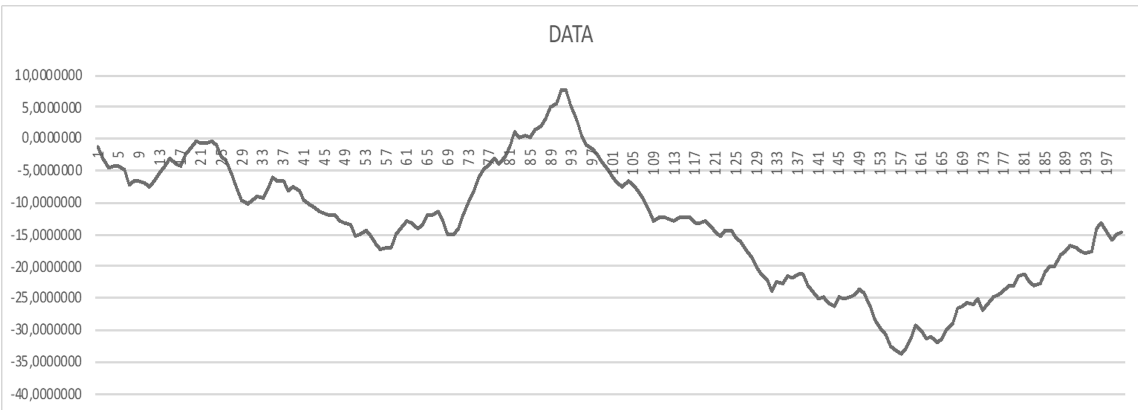


Figure 4.7. Plot Generated Artificial Data (Up & Down Trend, Non Seasonal/Cyclic)

Table 4.7. Forecasting accuracy of Generated Artificial Data (Up & Down Trend, Non Seasonal/Cyclic)

Data	Forecast						
	α-Sutte	ARIMA	Holt-Winter	NNETAR	Robust	Theta	Sutte ARIMA
-17.56	-18.44	-18.25	-18.28	-18.25	-18.26	-18.08	-18.35
-14.09	-17.71	-18.37	-18.55	-18.37	-18.45	-18.14	-18.04
-13.08	-13.08	-18.42	-18.82	-18.45	-18.61	-18.21	-15.75
-14.47	-11.57	-18.45	-19.09	-18.51	-18.74	-18.27	-15.01
-15.83	-13.60	-18.46	-19.36	-18.56	-18.84	-18.33	-16.03
-15.02	-16.46	-18.47	-19.62	-18.61	-18.93	-18.39	-17.46
-14.55	-15.72	-18.47	-19.89	-18.66	-18.99	-18.46	-17.09
Forecast accuracy							
MSE	**4.38**	13.90	19.56	14.45	16.07	12.90	5.16
MAPE	**11.77**	24.16	28.75	24.66	26.05	23.18	13.09
MAE	**1.75**	3.47	4.15	3.54	3.75	3.33	1.88

From the test results on data with an Up & Down trend, and non-seasonal/cyclic (table 4.7), it is found that the -Sutte method has a good level of accuracy based on the values of MSE, MAPE, and MAE.

8) The Generated Artificial Data (Trend, Non Seasonal/Cyclic)

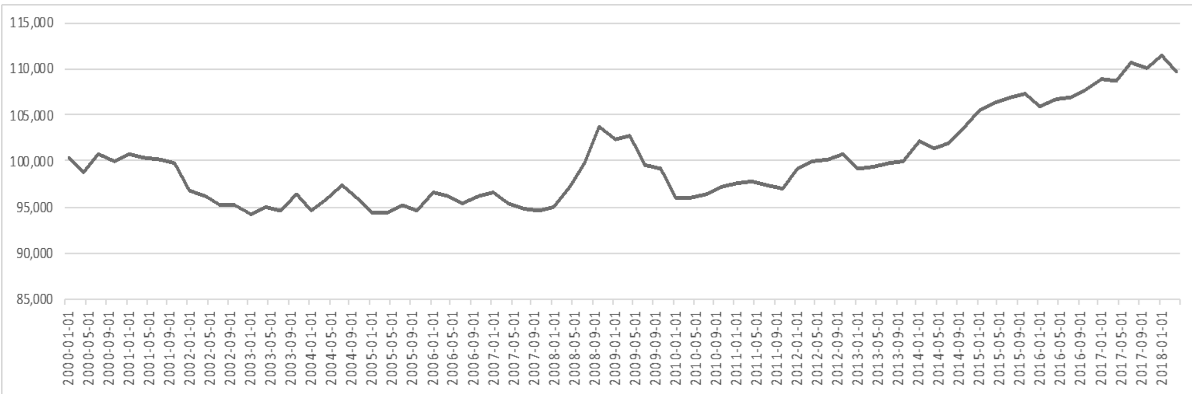


Figure 4.8. Plot Generated Artificial Data (Trend, Non Seasonal/Cyclic)

Table 4.8. Forecasting accuracy of Generated Artificial Data (Trend, Non Seasonal/Cyclic)

Data	Forecast						
	α-Sutte	ARIMA	Holt-Winter	NNETAR	Robust	Theta	Sutte ARIMA
107.77	106.82	106.97	107.36	106.98	105.44	107.02	106.90
108.90	108.35	106.97	107.76	107.00	104.97	107.08	107.66
108.75	109.65	106.97	108.15	107.01	104.60	107.14	108.31
110.81	109.35	106.97	108.54	107.02	104.28	107.19	108.16
110.05	111.83	106.97	108.94	107.02	103.99	107.25	109.40
111.53	110.44	106.97	109.33	107.03	103.73	107.30	108.71
109.78	112.46	106.97	109.73	107.04	103.49	107.36	109.71
Forecast accuracy							
MSE	2.24	8.64	**1.86**	8.37	31.10	7.30	2.56
MAPE	1.22	2.44	**1.01**	2.40	4.81	2.24	1.13
MAE	1.34	2.69	**1.11**	2.64	5.30	2.46	1.25

From the test results on trend and non-seasonal/cyclic data (table 4.8), it is found that the Holt-Winter method has a good level of accuracy based on the MSE, MAPE, and MAE values followed by the SutteARIMA, -Sutte, Theta, NNETAR, and ARIMA methods.

9) The Generated Artificial Data (Non Trend, Seasonal/Cyclic)

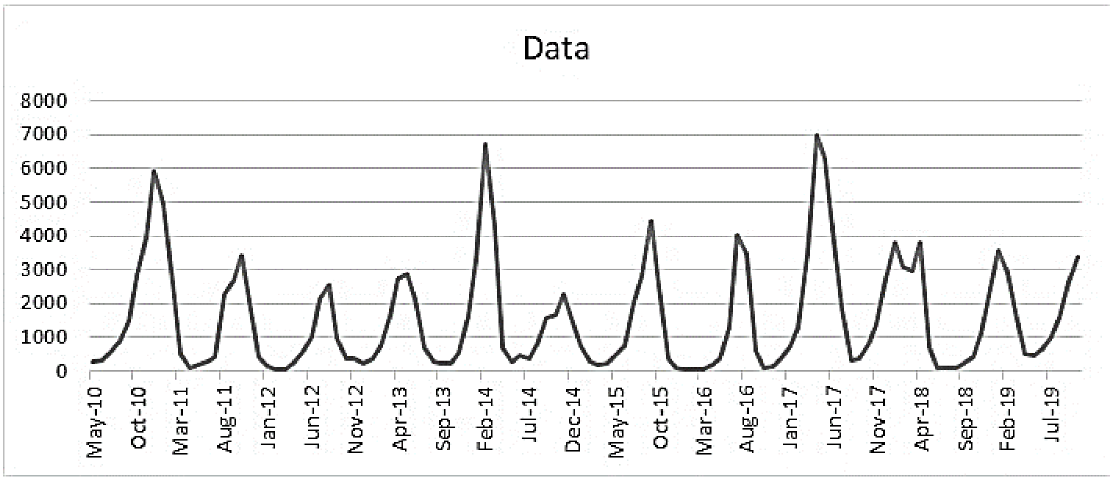


Figure 4.9. Plot Generated Artificial Data (Non Trend, Seasonal/Cyclic)

Table 4.9. Forecasting accuracy of Generated Artificial Data (Non Trend, Seasonal/Cyclic)

Data	Forecast						
	α-Sutte	ARIMA	Holt-Winter	NNETAR	Robust	Theta	Sutte ARIMA
529.00	1477.64	806.94	1589.00	569.18	2586.17	1539.18	1142.29
485.00	-155.48	669.12	1641.00	312.83	2743.88	1541.23	256.82
662,00	-21.42	860.10	1693.00	279.50	2901.59	1543.27	419.34
1000.00	544.01	1206.81	1745.00	346.03	3059.30	1545.31	875.41
1590.00	1189.65	1543.12	1797.00	519.79	3217.01	1547.36	1366.39
2657.00	2035.15	1762.06	1849.00	1358.00	3374.72	1549.40	1898.61
3396.00	3479.07	1831.12	1901.00	2882.72	3532.44	1551.44	2655.09
Forecast accuracy							
MSE	362712.07	492162.00	1001237.14	528778.82	3110260.73	1120152.88	**239528.67**
MAPE	73.04	**31.97**	108.06	42.51	218.89	99.30	39.50
MAE	547.68	481.95	928.86	590.19	1585.16	926.83	**418.80**

From the test results on Non-Trend and Seasonal/cyclic data (table 4.9), it is found that the SutteARIMA method has a good level of accuracy based on MSE and MAE values, while ARIMA on MAPE values.

10) The 1st Generated Artificial Data (Up & Down Trend, Seasonal/Cyclic)

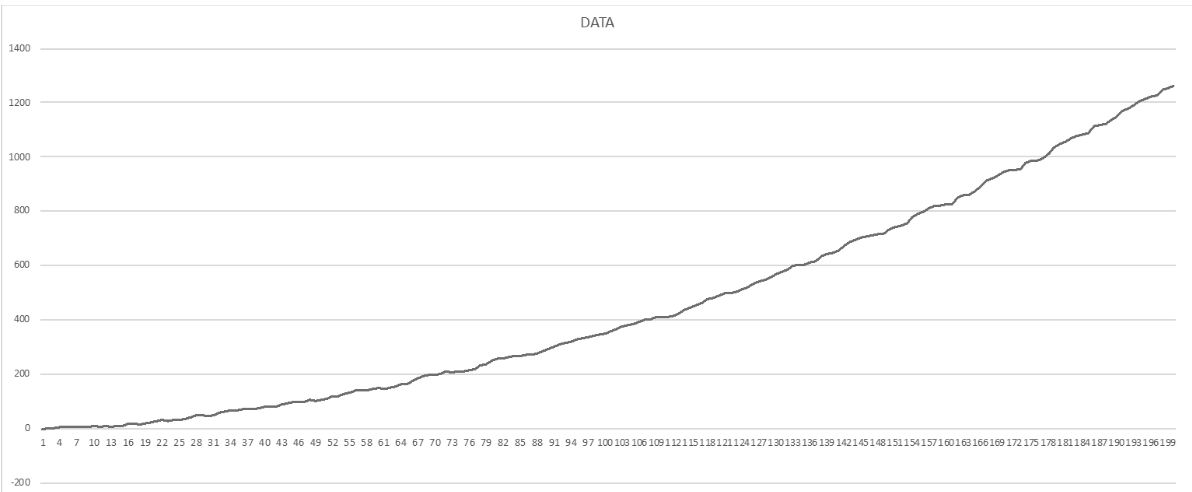


Figure 4.10. Plot 1st Generated Artificial Data (Up & Down Trend, Seasonal/Cyclic)

Table 4.10. Forecasting accuracy of 1st Generated Artificial Data (Up & Down Trend, Seasonal/Cyclic)

Data	Forecast						
	α-Sutte	ARIMA	Holt-Winter	NNETAR	Robust	Theta	Sutte ARIMA
1205.43	1206.40	1203.17	1203.12	1194.30	1203.32	1195.03	1204.79
1213.90	1217.20	1214.39	1214.28	1196.47	1214.70	1198.11	1215.79
1220.10	1225.30	1225.60	1225.45	1198.49	1226.07	1201.18	1225.45
1225.60	1229.52	1236.82	1236.61	1200.36	1237.45	1204.26	1233.17
1248.67	1232.34	1248.04	1247.77	1202.10	1248.83	1207.33	1240.19
1251.43	1260.34	1259.25	1258.94	1203.70	1260.21	1210.41	1259.80
1260.22	1261.95	1270.47	1270.10	1205.19	1271.58	1213.48	1266.21
Forecast accuracy							
MSE	57.63	46.87	44.31	1286.68	55.35	963.92	**38.26**
MAPE	0.47	0.44	0.43	2.59	0.47	2.25	**0.44**
MAE	5.77	5.45	5.33	32.11	5.86	27.94	**5.47**

11) The 2nd Generated Artificial Data (Up & Down Trend, Seasonal/Cyclic)

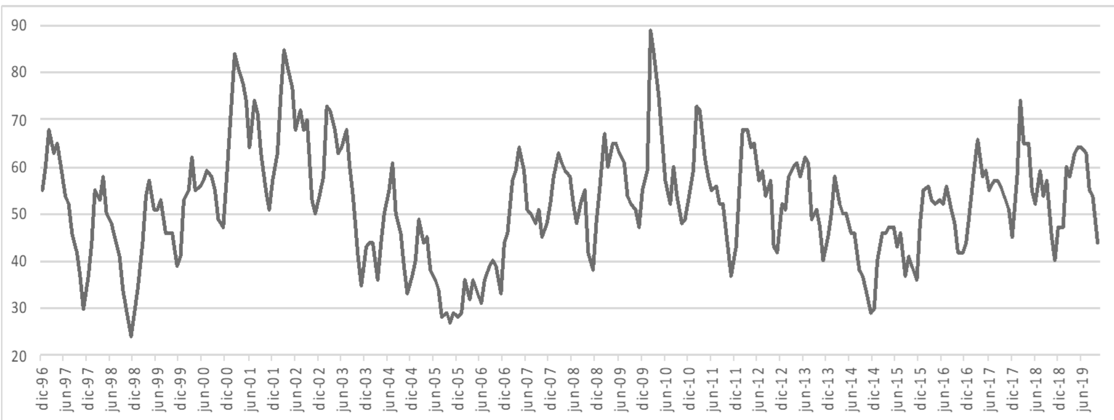


Figure 4.11. Plot 2nd Generated Artificial Data (Up & Down Trend, Seasonal/Cyclic)

Table 4.11. Forecasting accuracy of 2nd Generated Artificial Data (Up & Down Trend, Seasonal/Cyclic)

Data	Forecast						
	α-Sutte	ARIMA	Holt-Winter	NNETAR	Robust	Theta	Sutte ARIMA
63.00	62.20	56.12	58.20	59.44	58.00	58.00	59.16
64.00	68.94	54.83	58.41	54.25	58.00	58.00	61.88
64.00	65.42	52.48	58.61	54.72	58.00	58.00	58.95
63.00	66.07	53.20	58.81	59.97	58.00	57.99	59.64
55.00	63.01	53.14	59.02	52.92	58.00	57.99	58.07
54.00	52.18	53.09	59.22	57.72	58.00	57.99	52.64
44.00	50.85	53.04	59.42	48.52	58.00	57.99	51.95
Forecast accuracy							
MSE	21.55	63.75	54.60	34.52	49.00	48.97	**18.64**
MAPE	7.08	12.06	11.92	8.73	11.33	11.33	**6.97**
MAE	3.84	7.03	6.38	5.13	6.14	6.14	**3.82**

12) The Generated Artificial Data (Trend, Seasonal/Cyclic)

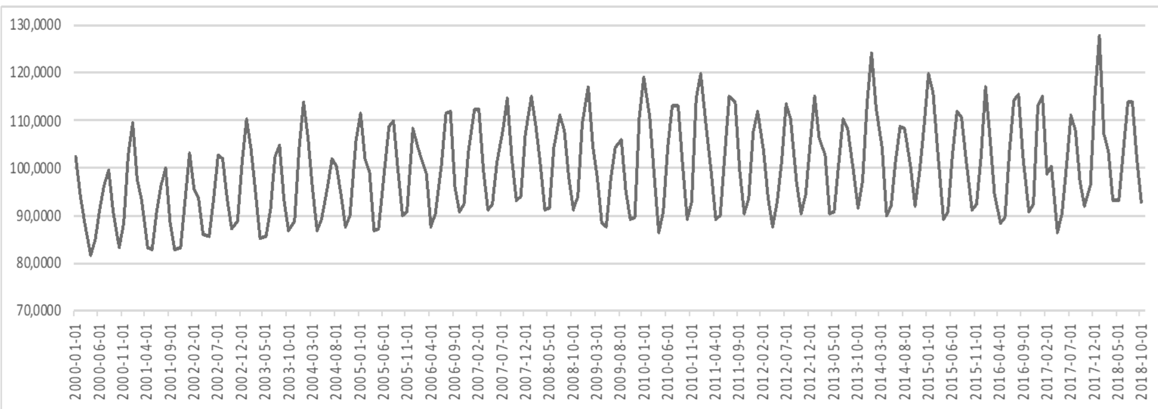


Figure 4.12. Plot Generated Artificial Data (Trend, Seasonal/Cyclic)

Table 4.12. Forecasting accuracy of Generated Artificial Data (Trend, Seasonal/Cyclic)

Data	Forecast						
	α-Sutte	ARIMA	Holt-Winter	NNETAR	Robust	Theta	Sutte ARIMA
93.20	100.33	97.04	99.78	94.20	103.57	103.59	98.68
93.31	82.51	101.99	96.16	93.07	103.76	103.62	92.25
103.67	88.95	115.86	92.55	101.52	103.94	103.64	102.41
113.78	104.12	119.32	88.93	104.49	104.12	103.67	111.72
113.87	120.99	112.25	85.32	98.38	104.30	103.69	116.62
101.87	121.05	102.33	81.70	93.68	104.49	103.72	111.69
92.79	101.65	97.80	78.09	90.37	104.67	103.74	99.73
Forecast accuracy							
MSE	139.29	42.49	318.61	57.84	78.52	77.65	**27.03**
MAPE	10.94	5.33	14.77	5.12	7.84	7.67	**4.23**
MAE	11.07	5.33	15.54	5.54	7.83	7.69	**4.20**

From the test results on data with Up & Down trends, and seasonal/cyclic (table 4.10 and table 4.11) as well as Trend, and seasonal/cyclic (table 4.12), it was found that the SutteARIMA method has a good level of accuracy based on the values of MSE, MAPE, and MAE.

The summary of the results of the data generated test is described in table 4.13.

Based on table 4.13, it can be seen that the SutteARIMA method is suitable for Non Trend, Trend, Non Seasonal, Seasonal data, and their combination. To see more about SutteARIMA, in the Empirical examples section, forecasting will be carried out on the generated data.

4.3. Methodology and data

4.3.1. Dataset

In the empirical examples section, forecasting will be carried out using 2 actual datasets (not artificial data), namely (1) National Currency to US Dollar Spot Exchange Rate for Indonesia and (2) Consumer Price Index: All Items for Indonesia.

Table 4.13. Summary of Test Results on Data Generate

Forecast Methods	Non trend + non seasonal	Up trend + non seasonal	Trend + non seasonal	Down trend & non seasonal	Up & down trend + non seasonal	Non trend + seasonal	Up & down trend + seasonal	Trend + seasonal
ARIMA	-	✓	-	-	-	✓	-	-
Holt-winter	✓	✓	✓	-	-	-	-	-
NNETAR	✓	-	-	-	-	-	-	-
Robust	-	✓	-	-	-	-	-	-
Theta	✓	✓	-	-	-	-	-	-
α-Sutte	✓	✓	-	-	✓	-	-	-
SutteARIMA	-	-	-	✓	✓	✓	✓	✓

The dataset of this study was obtained from the Federal Reserve Economic Data; (1) National Currency to US Dollar Spot Exchange Rate for Indonesia, US Dollar per National Currency Units, Monthly, Not Seasonally Adjusted (https://fred.stlouisfed.org/series/CCUSSP02IDM650N); and (2) Consumer Price Index: All Items for Indonesia, Index 2015=100, Monthly, Not Seasonally Adjusted (https://fred.stlouisfed.org/series/IDNCPIALLMINMEI).

Each dataset is divided into 2 parts, namely training data and testing data. Dataset I start from January 2018 to March 2021. Data training starts from January 2018 to August 2021, and data testing is carried out in the next 7 months from September 2021 to March 2021.

Dataset II start from January 2018 to April 2021. Data training starts from January 2018 to September 2021, and data testing is carried out in the next 7 months from October 2021 to April 2021.

4.3.2. Data analysis

The data in the study were analyzed using R Software. In forecasting, there are several R packages used, namely: SutteForecastR packages (Ahmar, 2017).

4.3.3. Forecast accuracy

A forecasting method is said to have a high level of accuracy if the value of forecast accuracy is the lowest value of all forecasting methods. In this research, forecast accuracy used is MSE and MAPE. The indicators are as follows (Ahmar, 2020).

- Mean absolute percentage error, MAPE

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right|$$

- Mean squared error, MSE

$$\text{MSE} = \frac{\sum_{t=1}^{n}\left(A_t - F_t\right)^2}{n}$$

Where:

A_t = Actual values at data time t.
F_t = Forecast values at data time t.

4.4. Empirical examples

4.4.1. Dataset I: National Currency to US Dollar Spot Exchange Rate for Indonesia

a. Descriptive statistics

Table 4.14. Descriptive Analysis of National Currency to US Dollar Spot Exchange Rate for Indonesia

US Dollar Exchange Rate	
Mean	14341
Standard Error	81
Median	14236
Standard Deviation	512
Sample Variance	262332
Kurtosis	5.3
Skewness	1.6
Range	2954
Minimum	13413
Maximum	16367
Sum	573631
Count	40
Mean	14341

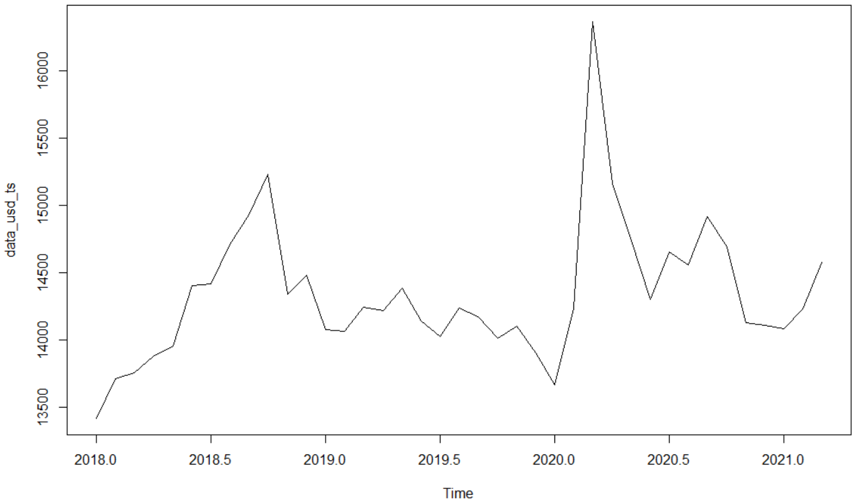


Figure 4.13. Time Series Plot of National Currency to US Dollar Spot Exchange Rate for Indonesia

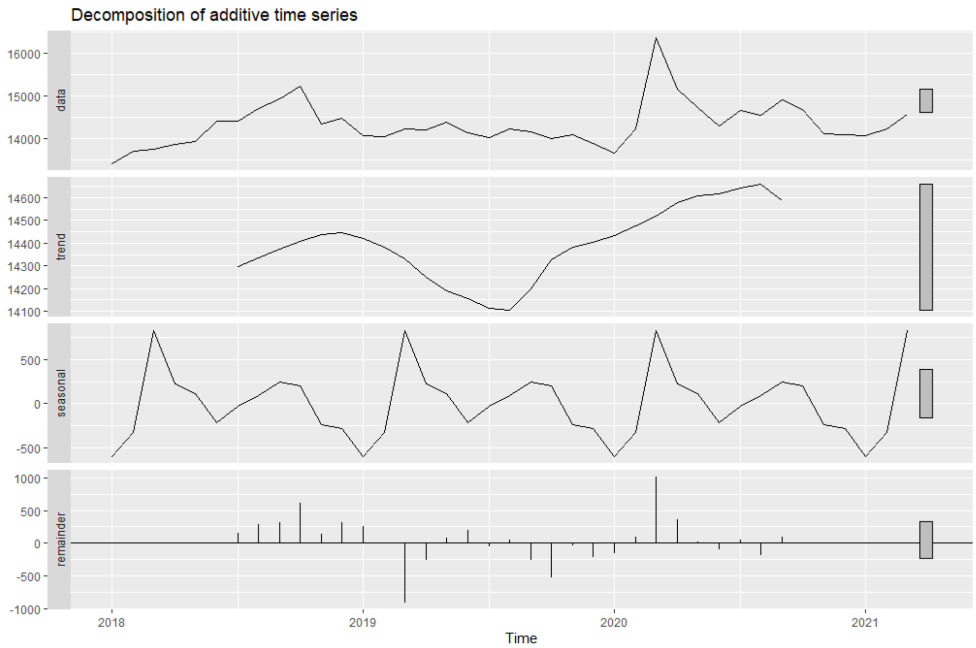


Figure 4.14. Time Series Decomposition Plot of National Currency to US Dollar Spot Exchange Rate for Indonesia

b. Estimation results

Forecasting results using R software are presented as follows.

$AutoARIMA
Series: al_mi_10
ARIMA(1,0,0) with non-zero mean

```
Coefficients:
         ar1        mean
      0.5313  14303.5007
s.e.  0.1535    169.9425

sigma^2 estimated as 230535:  log likelihood=-242.11
AIC=490.22   AICc=491.08   BIC=494.62
```

$HoltWinters
Holt-Winters exponential smoothing with trend and without seasonal component.

Call:
HoltWinters(x = al_mi_10, gamma = FALSE)

Smoothing parameters:
 alpha: 0.8232176
 beta: 0.08612287
 gamma: FALSE

Coefficients:
         [,1]
a 14571.61273
b   31.86022

$NNETAR
Series: al_mi_10
Model: NNAR(1,1)
Call: nnetar(y = al_mi_10)

Average of 20 networks, each of which is
a 1-1-1 network with 4 weights
options were - linear output units

sigma^2 estimated as 193901

$Robust_exponential_smoothing
ROBETS(M,A,N)

Call:
 robets(y = al_mi_10)

  Smoothing parameters:
    alpha = 0.6021
    beta = 0.1225

  Initial states:
    sigma = 0.0075
    l = 13191.6143
    b = 191.8143

  sigma: 0.0397

  robAIC robAICc   robBIC
461.8303 462.2441 464.7617

$Theta_Model
Theta

Call:
 forecast::ets(y = y, model = "ANN", opt.crit = "mse")

  Smoothing parameters:
    alpha = 0.6955

  Initial states:
    l = 13512.1703

  sigma: 526.637

     AIC   AICc   BIC
515.8950 516.7522 520.2923

Based on the results of the analysis presented, the forecasting model is obtained:

ARIMA(1,0,0) with non-zero mean with AR coefficient: 0.5313 or in mathematics:

$$Z_t = c + 0.5313Z_{t-1},$$
$$Z_t = \left(14303.5007 \times (1 - 0.5313)\right) + 0.5313Z_{t-1},$$
$$Z_t = 6704.051 + 0.5313Z_{t-1}.$$

The Holt-Winters formula of this data is Holt-Winters exponential smoothing with trend and without seasonal component or only using two parameters (α = 0.8232176 and β = 0.08612287).

The neural network time series, NNETAR, method is obtained by a forecasting model, namely NNAR (1,1) with 1 hidden layer. Robust Exponential Smoothing/ROBETS(M,A,N) with smoothing parameters: α = 0.6021 and β = 0.1225. And the Theta Model with smoothing parameter α = 0.6955 (simple exponential smoothing with additive errors).

c. Forecasting results

Forecasting results from National Currency to US Dollar Spot Exchange Rate for Indonesia can be seen in table 4.15 and figure 4.15.

Table 4.15. Forecasting Results from National Currency to US Dollar Spot Exchange Rate for Indonesia

Data	α-Sutte	ARIMA	Holt-Winter	NNETAR	Robust	Theta	Sutte ARIMA
14918	14498	14437	14603	14572	14627	14583	14467
14690	15126	14374	14635	14582	14653	14594	14750
14128	14705	14341	14667	14587	14678	14604	14523
14105	13992	14323	14699	14590	14704	14614	14158
14084	13838	14314	14731	14592	14729	14624	14076
14229	13886	14309	14763	14592	14754	14634	14097
14572	14263	14306	14795	14593	14780	14644	14285

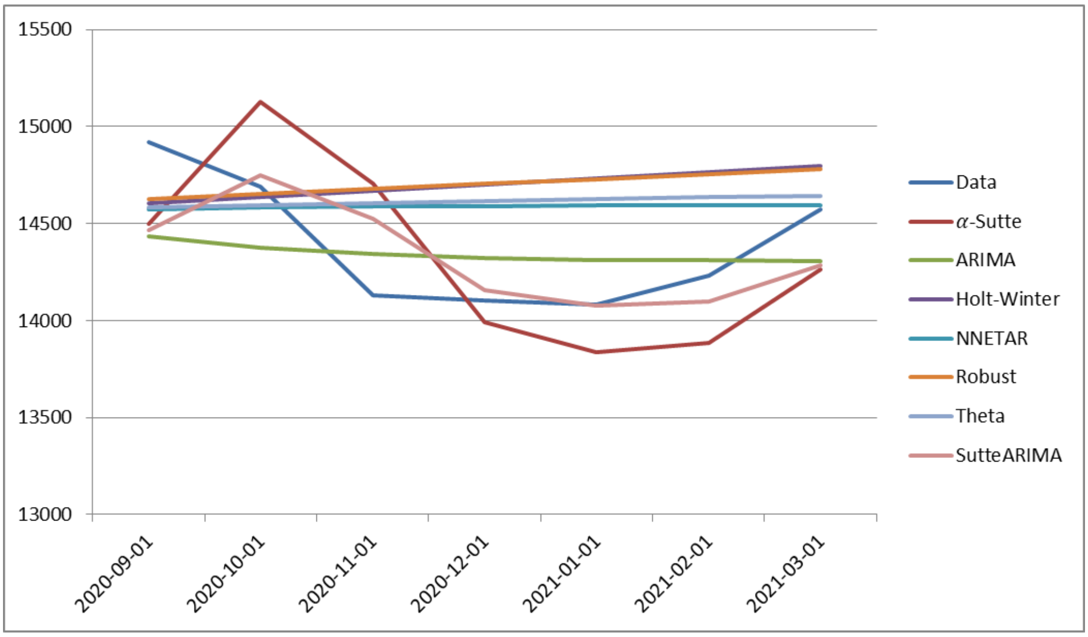


Figure 4.15. Plot of Forecasting Results from National Currency to US Dollar Spot Exchange Rate for Indonesia

4.4.2. Dataset II: Consumer Price Index: All Items for Indonesia

a. Descriptive statistics

Table 4.16. Descriptive Analysis of Consumer Price Index: All Items for Indonesia

Consumer Price Index: All Items for Indonesia	
Mean	114.2705041
Standard Error	0.412060341
Median	115.0083734
Standard Deviation	2.606098424
Sample Variance	6.791748995
Kurtosis	-1.281529054
Skewness	-0.347250871
Range	8.297332783
Minimum	109.6978672
Maximum	117.9952
Sum	4570.820164
Count	40

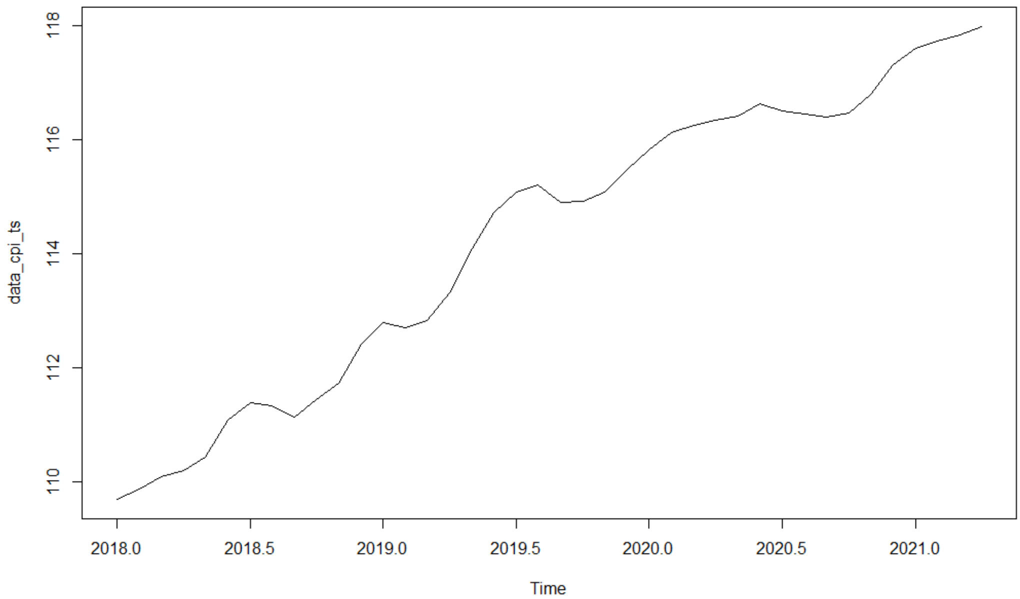


Figure 4.16. Time Series Plot of Consumer Price Index: All Items for Indonesia

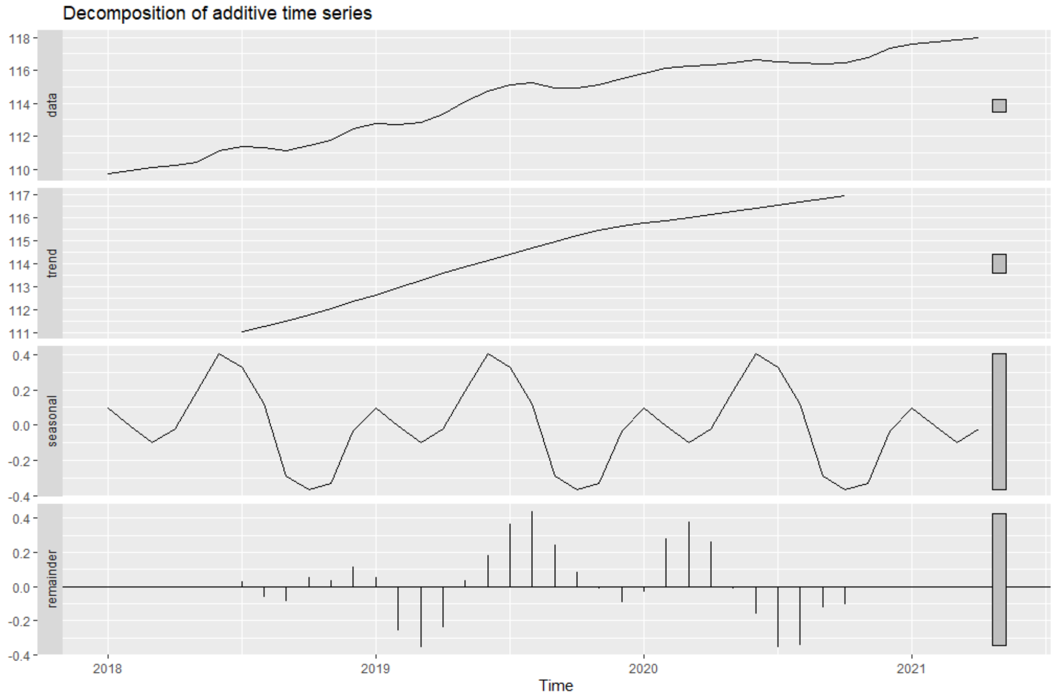


Figure 4.17. Time Series Decomposition Plot of Consumer Price Index: All Items for Indonesia

b. Estimation results

Forecasting results using R software are presented as follows.

$AutoARIMA
Series: al_mi_10
ARIMA(0,1,2) with drift

```
Coefficients:
         ma1     ma2   drift
      0.7518  0.3887  0.2028
s.e.  0.1904  0.1937  0.0769

sigma^2 estimated as 0.04792:  log likelihood=4.58
AIC=-1.17   AICc=0.32   BIC=4.7
```

$HoltWinters
Holt-Winters exponential smoothing with trend and without seasonal com ponent.

Call:
HoltWinters(x = al_mi_10, gamma = FALSE)

```
Smoothing parameters:
 alpha: 1
 beta: 1
 gamma: FALSE

Coefficients:
          [,1]
a 116.39664639
b  -0.05550627
```

$NNETAR

```
Series: al_mi_10
Model:  NNAR(1,1)
Call:   nnetar(y = al_mi_10)

Average of 20 networks, each of which is
a 1-1-1 network with 4 weights
options were - linear output units

sigma^2 estimated as 0.05404
```

$Robust_exponential_smoothing

```
ROBETS(A,Ad,N)

Call:
 robets(y = al_mi_10)

  Smoothing parameters:
    alpha = 0.3576
    beta  = 0.1162
    phi   = 0.9685

  Initial states:
    sigma = 0.101
    l = 109.4882
    b = 0.2001

  sigma:  0.3669

  robAIC  robAICc   robBIC
25.53115 26.35874 30.02067
```

$Theta_Model
Theta

Call:
forecast::ets(y = y, model = "ANN", opt.crit = "mse")

Smoothing parameters:
alpha = 0.9999

Initial states:
l = 109.6888

sigma: 0.3358

AIC	AICc	BIC
47.29099	48.11858	51.78051

Based on the results of the analysis presented, the forecasting model is obtained:

ARIMA(0,1,2) with drift with coefficient MA(1): 0.7518, MA(2): 0.3887, drift: 0.2028 or in mathematics:

$$Z_t = c + 0.5313 Z_{t-1},$$
$$Z_t = \left(14303.5007 \times (1 - 0.5313)\right) + 0.5313 Z_{t-1},$$
$$Z_t = 6704.051 + 0.5313 Z_{t-1}.$$

The Holt-Winters formula of this data is Holt-Winters exponential smoothing with trend and without seasonal component or only using two parameters ($\alpha = 1$ and $\beta = 1$).

The neural network time series, NNETAR, method is obtained by a forecasting model, namely NNAR (1,1) with 1 hidden layer. Robust Exponential Smoothing/ ROBETS(A,Ad,N) with smoothing parameters: $\alpha = 0.3576$, $\beta = 0.1162$, $\pi = 0.9685$. And the Theta Model with smoothing parameter $\alpha = 0.9999$ (simple exponential smoothing with additive errors).

c. Forecasting results

Forecasting results from the Consumer Price Index: All Items for Indonesia can be seen in table 4.17 and figure 4.18.

Table 4.17. Forecasting results from the Consumer Price Index: All Items for Indonesia

Data	α-Sutte	ARIMA	Holt-Winter	NNETAR	Robust	Theta	Sutte ARIMA
116.4744	116.3190	116.5101	116.3411	116.4170	116.6453	116.5171	116.4146
116.7963	116.4633	116.6706	116.2856	116.4324	116.6715	116.6376	116.5670
117.3181	116.9112	116.8735	116.2301	116.4441	116.6969	116.7580	116.8924
117.6178	117.6257	117.0763	116.1746	116.4530	116.7215	116.8785	117.3510
117.7399	117.9996	117.2791	116.1191	116.4597	116.7454	116.9990	117.6394
117.8398	118.0550	117.4820	116.0636	116.4647	116.7685	117.1194	117.7685
117.9952	118.0139	117.6848	116.0081	116.4685	116.7908	117.2399	117.8494

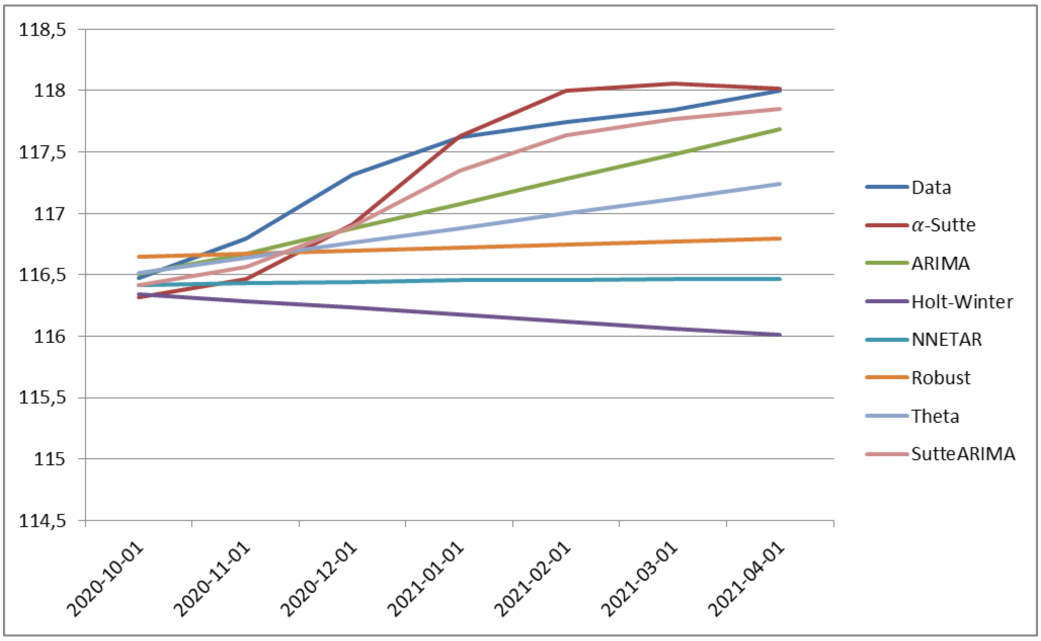


Figure 4.18. Plot of Forecast Results from Consumer Price Index: All Items for Indonesia

4.4.3. Assessing forecasting performances/forecast evaluation

a. National Currency to US Dollar Spot Exchange Rate for Indonesia

Based on table 4.18 and table 4.19, it can be seen that the SutteARIMA forecasting method has lower MSE and MAPE values than other forecasting methods. This indicates that SutteARIMA has a high level of accuracy and is suitable for use in forecasting the National Currency to US Dollar Spot Exchange Rate for Indonesia. This is reinforced by the time series decomposition (figure 4.14) which shows that the data are Up & Down Trend and Seasonal. In addition, this is in line with the opinion (Singh et al., 2021) which states that SutteARIMA is more suitable as compared to ARIMA method on stock market forecasting in USA; and the study (Ahmar et al., 2022b) which says that SutteARIMA is more suitable in predicting stock prices in Russia, India, and China than ARIMA and Holt-Winters.

b. Consumer Price Index: All Items for Indonesia

As only in forecasting for the National Currency to US Dollar Spot Exchange Rate for Indonesia, forecasting for the Consumer Price Index: All Items for Indonesia is also more suitable to use the SutteARIMA method, this is based on table 4.20 and table 4.21, it can be seen that the SutteARIMA forecasting method has a high level of accuracy. which is based on MSE and MAPE values which are smaller than other forecasting methods. This is reinforced by the time series decomposition (figure 4.16) which shows that the data are Up Trend and Seasonal. The study (Ahmar et al., 2022a) also explains that the MSE and MAPE values of SutteARIMA are smaller than ARIMA, Neural Networks Time Series (NNAR), and Holt-Winters in forecasting the Infant Mortality Rate in Indonesia.

4.5. Conclusion, Policy implication, and further research

Forecasting on linear and non-linear data for financial and economic data can be a guide and guide for decision makers. Based on forecasting trials using generated artificial data and actual data, it was found that the SutteARIMA method is more suitable than other methods (α-Sutte, ARIMA, Holt-Winter, NNETAR, Robust, and Theta). This is reinforced by the smaller MSE and MAPE values compared to other methods. From the research and test results on the generated artificial data, it is concluded that the SutteARIMA method is suitable for Non Trend, Trend, Non Seasonal, Seasonal data, and their combinations.

In addition, the findings of this study can be used by policy makers in making decisions and future impacts and of course can also be used as a tool in policy formulation. Through this research, research on forecasting of other financial and economic data can be tested and compared with other methods, such as time series linear model (TSLM), Long-Short Term Memory (LSTM), Recurrent Neural Networks (RNN), Support Vector Regression (SVR), ARIMAX, genetic programming (GP), or other methods.

Table 4.18. Mean squared error, MSE, value

Date	Data	α-Sutte	ARIMA	Holt-Winter	NNETAR	Robust	Theta	SutteARIMA
2020-09-01	14918	176408.40	231755.59	98929.12	119958.32	84500.68	111957.16	203139.50
2020-10-01	14690	190410.05	99723.32	2988.81	11765.74	1390.54	9298.74	3634.28
2020-11-01	14128	332386.84	45398.82	290725.86	210662.64	302621.01	226338.06	155867.04
2020-12-01	14105	12827.83	47724.77	352895.40	235186.20	358214.22	259009.74	2766.76
2021-01-01	14084	60393,06	52950.61	418492.55	257647.61	415908.91	291718.81	61.15
2021-02-01	14229	117882.36	6420.82	284910.41	132110.44	275950.60	164259.98	17319.88
2021-03-01	14572	95524.26	70495.56	49564.12	438.90	43143.44	5250.45	82535.54
MSE		**140833.26**	**79209.93**	**214072.32**	**138252.84**	**211675.63**	**152547.57**	**66474.88**

Table 4.19. Mean Absolute Percentage Error, MAPE, value

Date	Data	α-Sutte	ARIMA	Holt-Winter	NNETAR	Robust	Theta	SutteARIMA
2020-09-01	14918	2.82	3.23	2.11	2.32	1.95	2.24	3.02
2020-10-01	14690	2.93	2.12	0.37	0.73	0.25	0.65	0.40
2020-11-01	14128	3.86	1.43	3.61	3.08	3.69	3.19	2.65
2020-12-01	14105	0.76	1.46	3.98	3.25	4.01	3.41	0.35
2021-01-01	14084	1.65	1.54	4.34	3.40	4.32	3.62	0.05
2021-02-01	14229	2.30	0.54	3.58	2.44	3.52	2.72	0.88
2021-03-01	14572	2.07	1.78	1.49	0.14	1.39	0.49	1.93
MAPE		**2.34**	**1.73**	**2.78**	**2.19**	**2.73**	**2.33**	**1.33**

Table 4.20. Mean squared error, MSE, value

Date	Data	α-Sutte	ARIMA	Holt-Winter	NNETAR	Robust	Theta	SutteARIMA
2020-10-01	116.4744	0.0241	0.0013	0.0178	0.0033	0.0292	0.0018	0.0036
2020-11-01	116.7963	0.1109	0.0158	0.2608	0.1324	0.0156	0.0252	0.0526
2020-12-01	117.3181	0.1656	0.1977	1.1837	0.7639	0.3859	0.3137	0.1813
2021-01-01	117.6178	0.0001	0.2932	2.0828	1.3568	0.8034	0.5466	0.0712
2021-02-01	117.7399	0.0674	0.2123	2.6270	1.6389	0.9890	0.5489	0.0101
2021-03-01	117.8398	0.0463	0.1280	3.1549	1.8909	1.1477	0.5190	0.0051
2021-04-01	117.9952	0.0003	0.0963	3.9486	2.3308	1.4506	0.5705	0.0213
MSE		**0.0593**	**0.1350**	**1.8965**	**1.1596**	**0.6888**	**0.3608**	**0.0493**

Table 4.21. Mean Absolute Percentage Error, MAPE, value

Date	Data	α-Sutte	ARIMA	Holt-Winter	NNETAR	Robust	Theta	SutteARIMA
2020-10-01	116.4744	0.1334	0.0307	0.1144	0.0493	0.1467	0.0367	0.0514
2020-11-01	116.7963	0.2859	0.1079	0.4385	0.3124	0.1071	0.1363	0.1969
2020-12-01	117.3181	0.3493	0.3817	0.9341	0.7504	0.5333	0.4809	0.3655
2021-01-01	117.6178	0.0068	0.4649	1.2391	1.0000	0.7695	0.6347	0.2291
2021-02-01	117.7399	0.2230	0.3956	1.3916	1.0991	0.8538	0.6361	0.0863
2021-03-01	117.8398	0.1848	0.3072	1.5250	1.1806	0.9198	0.6185	0.0612
2021-04-01	117.9952	0.0161	0.2665	1.7060	1.3108	1.0340	0.6485	0.1252
MAPE		**0.1713**	**0.2792**	**1.0498**	**0.8147**	**0.6235**	**0.4559**	**0.1594**

Chapter 5
Conclusions

5.1. Main Conclusions

The main objective of this study was to develop a new forecasting method i.e. SutteARIMA method. SutteARIMA was developed by using a combination and/or weaknesses of some forecasting methods that already exist (α-Sutte Indicator and ARIMA). Based on the this main objective, it will be outlined into three specific objectives as follows: (1) development of new forecasting method (SutteARIMA) in the sector of finance; (2) forecasting of financial and actuarial data by using SutteARIMA method; (3) comparative study of financial data forecasting results.

The specific objectives of this research have been discussed and implemented so as to produce several general conclusions, namely: SutteARIMA method has a better level of accuracy for predicting economics, business, and actuarial data which has been discussed in Chapters 2, 3, and 4. The results of this conclusion are strengthened by obtaining the mean absolute percentage error, MAPE, and mean squared error, MSE, which is smaller when compared to other forecasting methods.

In Chapter 2, in the study it was obtained that the SutteARIMA method has a better level of accuracy compared to the ARIMA method seen from the mean absolute percentage error, MAPE, value which is then implemented in predicting the case of COVID-19 and the Stock Market in Spain.

In Chapter 3, the study showed that the accuracy level of SutteARIMA method (MAPE: 0.83% and MSE: 0.046) in predicting Infant Mortality rate in Indonesia was smaller than the other three forecasting methods, specifically the ARIMA (0.2.2) with a MAPE of 1.21 % and a MSE of 0.146; the NNAR with a MAPE of 7.95% and a MSE of 3.90; and the Holt-Winters with a MAPE of 1.03% and a MSE: of 0.083.

In Chapter 4, the study showed that the SutteARIMA method is more accurate than other models, based on MSE: 66474.88 / MAPE: 1.33% on National Currency to US Dollar Spot Exchange Rate for Indonesia and MSE: 0.0493 / MAPE: 0.1594% on Consumer Price data Index: All Items for Indonesia. In addition, from the results of research and test results on the generated artificial data, it is also found that SutteARIMA is suitable for Non Trend, Trend, Non Seasonal, Seasonal data, and their combination.

5.2. Limitations and Future Research

In addition to the resulting academic contributions, this research also has limitations that need to be considered in subsequent developments. In Chapter 2, which examines the implementation of SutteARIMA for the Stock Market in Spain, it may be necessary to conduct research on the Stock Market in various developed, developing and low-income countries. In Chapter 3 and Chapter 4, it is also necessary to compare the results using other forecasting methods using forecasting methods using the latest methods or other combination methods, such as the α-Sutte Indicator with Artificial Intelligence.

SutteARIMA is a short-term forecasting method, so this deficiency needs to be investigated further to be developed into a long-term forecasting method, in addition, it may be necessary to make a generalization formula for the previous data because SutteARIMA only uses 3 previous data for forecasting. In table 13 (Chapter 4) which discusses the summary of test results on data generate, it can be seen that SutteARIMA's accuracy in forecasting data (Non trend + non seasonal), (Up trend + non seasonal), and (Trend + non seasonal) is not as good as Holt-winter forecasting method; NNETAR; Robust; and α-Sutte, so it needs to be developed for this.

The SutteARIMA forecasting method is a forecasting method that is still new, so it requires input and improvements so that it can become a reliable forecasting method in various fields. Based on this, SutteARIMA will continue to carry out a continuous improvement process in order to obtain a good and feasible method.

References

Abonazel, M. R., & Abd-Elftah, A. I. (2019). Forecasting Egyptian GDP Using ARIMA Models. *Reports on Economics and Finance*, 5(1), 35–47. https://doi.org/10.12988/ref.2019.81023

Adeyinka, D. A., & Muhajarine, N. (2020). Time series prediction of under-five mortality rates for Nigeria: comparative analysis of artificial neural networks, Holt-Winters exponential smoothing and autoregressive integrated moving average models. *BMC Medical Research Methodology*, *20*(1), 292. https://doi.org/10.1186/s12874-020-01159-9

Ahmar, A. S. (2017). *sutteForecastR: Forecasting Data using Alpha-Sutte Indicator*. R package version 0.1, https://cran.r-project.org/web/packages/sutteForecastR/index.html

Ahmar, A. S. (2018). A Comparison of α-Sutte Indicator and ARIMA Methods in Renewable Energy Forecasting in Indonesia. *International Journal of Engineering and Technology(UAE)*, *7*(1.6), 20–22. www.sciencepubco.com/index.php/IJET

Ahmar, A. S. (2019). *Reliability Test of SutteARIMA to Forecast Artificial Data*. https://doi.org/10.31219/osf.io/9zn7v

Ahmar, A. S. (2020). Forecast Error Calculation with Mean Squared Error (MSE) and Mean Absolute Percentage Error (MAPE). *JINAV: Journal of Information and Visualization*, 1(2). https://doi.org/10.35877/454RI.jinav303

Ahmar, A. S., & Boj, E. (2020a). SutteARIMA: Short-term forecasting method, a case: Covid-19 and stock market in Spain. *Science of The Total Environment*, 138883. https://doi.org/10.1016/j.scitotenv.2020.138883

Ahmar, A. S., & Boj, E. (2020b). The date predicted 200.000 cases of COVID-19 in Spain. *Journal of Applied Science, Engineering, Technology, and Education*, 2(2). https://doi.org/10.35877/454RI.asci22102

Ahmar, A. S., Boj, E., A. El Safty, M., AlZahrani, S., & El-Khawaga, H. (2022). SutteARIMA: A Novel Method for Forecasting the Infant Mortality Rate in Indonesia. *Computers, Materials & Continua*, 70(3), 6007–6022. https://doi.org/10.32604/cmc.2022.021382

Ahmar, A. S., Guritno, S., & Abdurakhman. (2013). *Pendeteksian dan Pengoreksian Data yang Mengandung Additive Outlier (AO) pada Model ARIMA(p,d,q)*. Gadjah Mada University.

Ahmar, A. S., Rahman, A., Arifin, A. N. M., & Ahmar, A. A. (2017). Predicting movement of stock of "Y" using sutte indicator. *Cogent Economics and Finance*, 5(1). https://doi.org/10.1080/23322039.2017.1347123

Ahmar, A. S., Rahman, A., & Mulbar, U. (2018). α-Sutte Indicator: A new method for time series forecasting. *Journal of Physics: Conference Series*, 1040(1). https://doi.org/10.1088/1742-6596/1040/1/012018

Ahmar, A. S., Singh, P. K., Van Thanh, N., Viet Tinh, N., & Minh Hieu, V. (2022). Prediction of BRIC Stock Price Using ARIMA, SutteARIMA, and Holt-Winters. *Computers, Materials & Continua*, 70(1), 523–534. https://doi.org/10.32604/cmc.2022.017068

Aiolfi, M., & Timmermann, A. (2006). Persistence in forecasting performance and conditional combination strategies. *Journal of Econometrics*, 135(1–2), 31–53.

Anokye, R., Acheampong, E., Owusu, I., & Isaac Obeng, E. (2018). Time series analysis of malaria in Kumasi: Using ARIMA models to forecast future incidence. *Cogent Social Sciences*, 4(1). https://doi.org/10.1080/23311886.2018.1461544

Bakri, R., Data, U., & Saputra, A. (2019). Marketing Research : The Application of Auto Sales Forecasting Software to Optimize Product Marketing Strategies. *Journal of Applied Science, Engineering, Technology, and Education*, 1(1), 6–12. https://doi.org/10.35877/454ri.asci1124

Benvenuto, D., Giovanetti, M., Vassallo, L., Angeletti, S., & Ciccozzi, M. (2020). Application of the ARIMA model on the COVID-2019 epidemic dataset. *Data in Brief*, 29, 105340. https://doi.org/10.1016/j.dib.2020.105340

Becker, R., & Clements, A. E. (2008). Are combination forecasts of S&P 500 volatility statistically superior? *International Journal of Forecasting*, 24(1), 122–133.

Billah, B., King, M. L., Snyder, R. D., & Koehler, A. B. (2006). Exponential smoothing model selection for forecasting. *International Journal of Forecasting*, 22(2), 239–247. https://doi.org/10.1016/j.ijforecast.2005.08.002

Brockwel, P. J., & Davis, R. A. (2006). *Time series : Theory and Methods* (2nd ed.). Springer-Verlag.

Brockwell, P. J., & Davis, R. A. (2016). *Introduction to time series and forecasting* (2nd ed.). Springer.

Chakrabortya, I., & Maity, P. (2020). COVID-19 outbreak: Migration, effects on society, global environment and prevention. *Science of The Total Environment*, 728(138882). https://doi.org/10.1016/j.scitotenv.2020.138882

Clemen, R. T. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting*, 5(4), 559–583.

Crevits, R., Bergmeir, C., & Hyndman, R. (2018). *robets: Forecasting Time Series with Robust Exponential Smoothing*. R package version 1.4, https://cran.r-project.org/web/packages/robets/index.html

Dixon, H. (2020). The measurement of inflation during the lockdown: a trial calculation. *National Institute of Economic and Social Research*. Retrieved from https://www.niesr.ac.uk/publications/measurement-inflation-during-lockdown-trial-calculation

Edwards, J. (2019, August). What is predictive analytics? Transforming data into future insights. *CIO*. https://www.cio.com/article/3273114/what-is-predictive-analytics-transforming-data-into-future-insights.html

El Safty, M. A. (2021). Modeling uncertainty knowledge of the topological methods. *Poincare Journal of Analysis & Applications*, 8(1), 89–102. https://doi.org/10.33786/pjaa.2021.v08i01(i).009

El Safty, M. A., & Al Zahrani, S. (2021). Topological Modeling for Symptom Reduction of Coronavirus. *Punjab University Journal of Mathematics*, 53(3), 47–59.

Emamgholipour, S., Arab, M., & Mohajerzadeh, Z. (2017). Life insurance demand: middle East and North Africa. *International Journal of Social Economics*, 44(4), 521–529. https://doi.org/10.1108/IJSE-04-2015-0106

Fanelli, D., & Piazza, F. (2020). Analysis and forecast of COVID-19 spreading in China, Italy and France. *Chaos, Solitons & Fractals*, 134, 109761. https://doi.org/10.1016/j.chaos.2020.109761

France24. (2020). *Spain announces lockdown after reporting 1,500 new coronavirus cases in a day*. https://www.france24.com/en/20200314-spain-to-announce-lockdown-after-reporting-1-500-new-coronavirus-cases-in-a-day

Gokcan, S. (2000). Forecasting volatility of emerging stock markets: linear versus non-linear GARCH models. *Journal of Forecasting*, 19(6), 499–504. https://doi.org/10.1002/1099-131X(200011)19:6<499::AID-FOR745>3.0.CO;2-P

Goodwin, P. (2014). Using Naave Forecasts to Assess Limits to Forecast Accuracy and the Quality of Fit of Forecasts to Time Series Data. In *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2515072

Grover, L. K., & Mehra, R. (2008). The Lure of Statistics in Data Mining. *Journal of Statistics Education*, 16(1). https://doi.org/10.1080/10691898.2008.11889552

Haiss, P., & Sümegi, K. (2008). The relationship between insurance and economic growth in Europe: a theoretical and empirical analysis. *Empirica*, 35(4), 405–431. https://doi.org/10.1007/s10663-008-9075-2

Hansen, P. R., Lunde, A., & Nason, J. M. (2003). Choosing the best volatility models: the model confidence set approach. *Oxford Bulletin of Economics and Statistics*, 65, 839–861.

Hibon, M., & Evgeniou, T. (2005). To combine or not to combine: selecting among forecasts and their combinations. *International Journal of Forecasting*, 21(1), 15–24.

Hussein, M. A. (1991). Analysis and forecasting infant mortality rate (IMR) in Egypt until year 2000. *The Egyptian Population and Family Planning Review*, 25(2), 32–46.

Hyndman, R., Athanasopoulos, G., Bergmeir, C., Caceres, G., Chhay, L., O'Hara-Wild, M., Petropoulos, F., Razbash, S., Wang, E., & Yasmeen, F. (2021). *forecast: Forecasting functions for time series and linear models*. R package version 8.15, https://cran.r-project.org/web/packages/forecast/index.html

IMF (2018). *World Economic Outlook (WEO) World Economic Outlook, April 2018 Cyclical Upswing, Structural Change*. International Monetary Fund (IMF).

Jayanti, K. D., Wijaya, E. W., Bisono, E. F., Nurkhalim, R. F., Susilowati, I., Setyawan, I., & Putra, B. R. (2020). Proyeksi Angka Kematian Bayi di Rumah Sakit X Kabupaten Kediri dengan Single Exponential Smoothing. *Jurnal Berkala Kesehatan*, *6*(2), 50. https://doi.org/10.20527/jbk.v6i2.8925

Kim, S., & Kim, H. (2016). A new metric of absolute percentage error for intermittent demand forecasts. *International Journal of Forecasting*, 32(3), 669–679. https://doi.org/10.1016/j.ijforecast.2015.12.003

Koczkodaj, W. W., Mansournia, M. A., Pedrycz, W., Wolny-Dominiak, A., Zabrodskii, P. F., Strzaška, D., Armstrong, T., Zolfaghari, A. H., Debski, M., & Mazurek, J. (2020). 1,000,000 cases of COVID-19 outside of China: The date predicted by a simple heuristic. *Global Epidemiology*, 100023. https://doi.org/10.1016/j.gloepi.2020.100023

Kumar, S. V., & Vanajakshi, L. (2015). Short-term traffic flow prediction using seasonal ARIMA model with limited input data. *European Transport Research Review*, 7(3), 21. https://doi.org/10.1007/s12544-015-0170-8

Kurniasih, N., Ahmar, A. S., Hidayat, D. R., Agustin, H., & Rizal, E. (2018). Forecasting Infant Mortality Rate for China: A Comparison Between α-Sutte Indicator, ARIMA, and Holt-Winters. *Journal of Physics: Conference Series*, 1028(1), 012195.

Lestari, S., Ahmar, A. S., & Ruliana, R. (2020). Eksplorasi Metode Triple Exponential Smoothing Pada Peramalan Jumlah Penggunaan Air Bersih di PDAM Kota Makassar. *VARIANSI: Journal of Statistics and Its Application on Teaching and Research*, 2(3), 128–146. https://doi.org/10.35580/variansiunm14641

Lin, Y., Chen, M., Chen, G., Wu, X., & Lin, T. (2015). Application of an autoregressive integrated moving average model for predicting injury mortality in Xiamen, China. *BMJ Open*, 5(12), e008491.

https://doi.org/10.1136/bmjopen-2015-008491

Liu, Q., Liu, X., Jiang, B., & Yang, W. (2011). Forecasting incidence of hemorrhagic fever with renal syndrome in China using ARIMA model. *BMC Infectious Diseases*, *11*(1), 218. https://doi.org/10.1186/1471-2334-11-218

Maechler, M. (2020). *fracdiff: Fractionally Differenced ARIMA aka ARFIMA(P,d,q) Models.* R package version 1.5-1, https://cran.r-project.org/web/packages/fracdiff/index.html

Mahmudah, U. (2017). Predicting unemployment rates in Indonesia. *Economic Journal of Emerging Markets*, 9(1), 20. https://doi.org/10.20885/ejem.vol9.iss1.art3

McMurtry, A. (2020). *COVID-19: Spain unveils lifeline amid stock market bust.* https://www.aa.com.tr/en/latest-on-coronavirus-outbreak/covid-19-spain-unveils-lifeline-amid-stock-market-bust/1764097

Mishra, A. K., Sahanaa, C., & Manikandan, M. (2019). Forecasting Indian infant mortality rate: An application of autoregressive integrated moving average model. *Journal of Family & Community Medicine*, 26(2), 123. https://doi.org/10.4103/jfcm.JFCM_51_18

Moghaddam, A. H., Moghaddam, M. H., & Esfandyari, M. (2016). Stock market index prediction using artificial neural network. *Journal of Economics, Finance and Administrative Science*, 21(41), 89–93. https://doi.org/10.1016/j.jefas.2016.07.002

Molina, L. L., Angón, E., García, A., Moralejo, R. H., Caballero-Villalobos, J., & Perea, J. (2018). Time series analysis of bovine venereal diseases in La Pampa, Argentina. *PLOS ONE*, 13(8), e0201739. https://doi.org/10.1371/journal.pone.0201739

Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015). *Introduction to time series analysis and forecasting*. John Wiley & Sons.

Nyoni, T., & Nathaniel, S. P. (2018). *Modeling rates of inflation in Nigeria: an application of ARMA, ARIMA and GARCH models.*

Ojugo, A. A., & Yoro, R. E. (2020). Predicting Futures Price And Contract Portfolios Using The ARIMA Model: A Case of Nigeria's Bonny Light and Forcados. *Quantitative Economics and Management Studies*, 1(4), 237–248. https://doi.org/10.35877/454RI.qems139

Patowary, A. N., Dutta, B., Barman, M. P., & Gadde, S. R. (2018). Accidental Deaths in India: Forecasting with ARIMA Model. *Environment and Ecology*, 36(3), 761–766.

Pauwels, L., & Vasnev, A. (2014). Forecast combination for U.S. recessions with real-time data. *The North American Journal of Economics and Finance*, 28, 138–

148. https://doi.org/10.1016/j.najef.2014.02.005

Pemberton, J. (1989). Forecast accuracy of nonlinear time series models'. *Technical Report*, 77.

Qeadan, F., Honda, T., Gren, L. H., Dailey-Provost, J., Benson, L. S., VanDerslice, J. A., Porucznik, C. A., Waters, A. B., Lacey, S., & Shoaf, K. (2020). Naive Forecast for COVID-19 in Utah Based on the South Korea and Italy Models-the Fluctuation between Two Extremes. *International Journal of Environmental Research and Public Health*, 17(8), 2750. https://doi.org/10.3390/ijerph17082750

Rajia, S., Sabiruzzaman, M., Islam, M. K., Hossain, M. G., & Lestrel, P. E. (2019). Trends and future of maternal and child health in Bangladesh. *PloS One*, 14(3), e0211875. https://doi.org/10.1371/journal.pone.0211875

Ranapurwala, S. I., Cavanaugh, J. E., Young, T., Wu, H., Peek-Asa, C., & Ramirez, M. R. (2019). Public health application of predictive modeling: an example from farm vehicle crashes. *Injury Epidemiology*, 6(1), 31. https://doi.org/10.1186/s40621-019-0208-9

Ranjbar, M., & Khaledian, M. (2014). Using Arima Time Series Model in Forecasting the Trend of Changes in Qualitative Parameters of Sefidrud River. *International Research Journal of Applied and Basic Sciences*, 8(3), 346–351. www.irjabs.com

Rizal, J. G. (2020). *Pandemi Covid-19, Apa Saja Dampak pada Sektor Ketenagakerjaan Indonesia?* Kompas.Com. https://www.kompas.com/tren/read/2020/08/11/102500165/pandemi-covid-19-apa-saja-dampak-pada-sektor-ketenagakerjaan-indonesia-?page=all

Roosa, K., Lee, Y., Luo, R., Kirpich, A., Rothenberg, R., Hyman, J. M., Yan, P., & Chowell, G. (2020). Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020. *Infectious Disease Modelling*, 5, 256–263. https://doi.org/10.1016/j.idm.2020.02.002

Shaub, D. & Ellis, P. (2020). *forecastHybrid: Convenient Functions for Ensemble Time Series Forecasts*. R package version 5.0.19, https://cran.r-project.org/web/packages/forecastHybrid/index.html

Singh, P. K., Chouhan, A., Bhatt, R. K., Kiran, R., & Ahmar, A. S. (2021). Implementation of the SutteARIMA method to predict short-term cases of stock market and COVID-19 pandemic in USA. *Quality & Quantity*. https://doi.org/10.1007/s11135-021-01207-6

Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, 70, 263–286. https://doi.org/10.1016/j.jbusres.2016.08.001

Sutiksno, D. U., Ahmar, A. S., Kurniasih, N., Susanto, E., & Leiwakabessy, A. (2018). Forecasting Historical Data of Bitcoin using ARIMA and α-Sutte Indicator. *Journal of Physics: Conference Series*, 1028(1), 012194.

Torri, T., & Vaupel, J. W. (2012). Forecasting life expectancy in an international context. *International Journal of Forecasting*, 28(2), 519–531. https://doi.org/10.1016/j.ijforecast.2011.01.009

Tratar, L. F., & Ahmar, A. S. (2019). The Comparison of Holt-Winters Methods and α-Sutte Indicator in Forecasting the Foreign Visitor Arrivals in Indonesia, Malaysia, and Japan. In L. Z. Stirn, M. K. Borštar, J. Žerovnik, S. Drobne, & J. Povh (Eds.), *Proceedings of the 15th International Symposium on Operational Research* (pp. 587–592). Slovenian Society Informatika - Section for Operational Research. http://fgg-web.fgg.uni-lj.si/~/sdrobne/sor/SOR%2719 - Proceedings.pdf

United Nations. (2020). COVID-19: impact could cause equivalent of 195 million job losses, says ILO chief.

Wang, C.-C. (2011). A comparison study between fuzzy time series model and ARIMA model for forecasting Taiwan export. *Expert Systems with Applications*, 38(8), 9296–9304. https://doi.org/10.1016/j.eswa.2011.01.015

Wang, Y., Shen, Z., & Jiang, Y. (2018). Comparison of ARIMA and GM(1,1) models for prediction of hepatitis B in China. *PLOS ONE*, 13(9), e0201987. https://doi.org/10.1371/journal.pone.0201987

Wei, W. W. S. (2006). *Time Series Analysis : Univariate and Multivariate Methods* (2nd ed.). Pearson Addison Wesley.

WHO. (2020). *Situation report - 77 Coronavirus disease 2019 (COVID-19*. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/

Winkler, R. L., & Makridakis, S. (1983). The combination of forecasts. *Journal of the Royal Statistical Society: Series A (General)*, 146(2), 150–157.

Worldometer. (2020). *Spain Coronavirus*. https://www.worldometers.info/coronavirus/country/spain/

Wuryandari, M. D., & Afrianto, I. (2012). Perbandingan Metode Jaringan Syaraf Tiruan Backpropagation Dan Learning Vector Quantization Pada Pengenalan Wajah. *Jurnal Komputer Dan Informatika (Komputa)*, 1(1).

Yang, Y., Peng, F., Wang, R., Guan, K., Jiang, T., Xu, G., Sun, J., & Chang, C. (2020). The deadly coronaviruses: The 2003 SARS pandemic and the 2020 novel coronavirus epidemic in China. *Journal of Autoimmunity*, 102434. https://doi.org/10.1016/j.jaut.2020.102434

Yao, X., Crook, J., & Andreeva, G. (2015). Support vector regression for loss given default modelling. *European Journal of Operational Research*, 240(2), 528–538. https://doi.org/10.1016/j.ejor.2014.06.043

Zhang, X., Zhang, T., Young, A. A., & Li, X. (2014). Applications and Comparisons of Four Time Series Models in Epidemiological Surveillance Data. *PLoS ONE*, 9(2), e88075. https://doi.org/10.1371/journal.pone.0088075