AN ANALYSIS AND FORECAST OF THE MOTOR PRIVATE INSURANCE CLAIM AMOUNTS IN KENYA USING THE ARIMA MODEL

Ong'ale Edith Lorna Alwang'a

126184

Submitted in partial fulfilment of the requirements for the Degree of Bachelor of Business Science in Actuarial Science at Strathmore University

Strathmore Institute of Mathematical Sciences

Strathmore University

Nairobi, Kenya

January 2025

DECLARATION

I declare that this work has not been previously submitted and approved for the award of a degree by this or any other University. To the best of my knowledge and belief, the Research Project contains no material previously published or written by another person except where due reference is made in the Research Project itself.

Ong'ale Edith Lorna Alwang'a [*Name of Candidate*]

[*Signature*]

Monday 20/1/2025 [*Date*]

This Research Project has been submitted for examination with my approval as the Supervisor.

Dr Sebastian Kaweto [*Name of Candidate*]

[*Signature*]

Wednesday 22/1/2025 [*Date*]

Strathmore Institute of Mathematical Sciences
Strathmore University

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ACF	Autocorrelation Function
ADF	Augmented Dickey-Fuller
AIC	Akaike Information Criterion
AKI	Association of Kenya Insurers
ARIMA	Autoregressive Integrated Moving Average
BIC	Bayesian Information Criterion
BF	Bornhuetter-Ferguson
CLM	Chain Ladder Method
ELR	Expected Loss Ratio
IBNR	Incurred But Not Reported
IRA	Insurance Regulatory Authority
PACF	Partial Autocorrelation Function
SARIMA	Seasonal Autoregressive Integrated Moving Average
TPFT	Third Party Fire and Theft
TPO	Third Party Only

ABSTRACT

General insurance companies that offer motor private insurance packages face the need to hold enough reserves in order to meet any future claims liabilities. This involves carrying out a forecast of the future possible claims experiences to get an idea of the expected outflow of money from the business. In an attempt to achieve this, actuaries in Kenya turn to the commonly used actuarial forecasting techniques, which are the Chain Ladder Method and the Bornhuetter-Ferguson method. In as much as these methods are deemed simplistic and straightforward, they are not flexible and may be prone to distortion if at all the claims reporting pattern changes. By utilizing time series analysis and forecasting techniques, particularly the ARIMA model, the study seeks to offer a more flexible and accurate approach to forecasting insurance claims. This study explores the use of the ARIMA model as one of the methods that can be used to forecast motor private insurance claims in Kenya specifically focusing on the motor private claims amount data for the top five general insurance companies by market share: Old Mutual General Insurance, APA General Insurance, GA Insurance, CIC General Insurance, and Britam General Insurance. Using secondary data from the Insurance Regulatory Authority (IRA) spanning 2013 to 2022, the study gives the descriptive characteristics of the data and identifies an optimal order of the ARIMA model for each of the five companies: UAP with ARIMA(1,1,1), APA General Insurance with ARIMA(1,1,2), GA with ARIMA(0,1,1), CIC General Insurance with ARIMA(0,1,1), and lastly Britam General Insurance with ARIMA(0,1,1).The test for the accuracy using the Ljung-Box test after the forecast is generated for the period 2023 to 2027 reveals that the optimal order of the ARIMA model for each of the companies is indeed a good fit for the respective data. This suggests that actuaries can adopt time series analysis and forecasting techniques, most especially the ARIMA models, when performing forecasts of claim amounts within the actuarial space.

CHAPTER ONE: INTRODUCTION

1.1. Background of the study

Claims form part of the main cash outflow by any insurance company, therefore generating an accurate forecast for the same is of the outmost importance given the uncertainties surrounding future claim experiences. Claims forecasting is important for insurance companies for a number of reasons that dictate the stability and continuity of operations.

Accurate and prudent claims forecasting ensures that the premium amount charged by the insurer is sufficient to cover the expected total claims. This in turn allows for adequate reserving of the premiums collected and informed investment decisions in attempts to meet the regulatory standard for reserving and to maintain solvency. A clear view of the possible future claims experience also gives a clear indication of the level of risk that the insurer is exposed to, and this could help the insurer to decide whether to take on reinsurance or review existing reinsurance contracts as a way of spreading out the risk that they are exposed to. In the long run, accurate and clear forecasts ensures that the company is stable and the overall satisfaction of the shareholders and the (Poufinuas et. al, 2023).

This study focuses on the claim experiences under the general insurance business, specifically the motor private insurance business in Kenya. There are about 35 general insurance companies in Kenya that offer motor private insurance. Motor private insurance policy is designed to protect individuals who own private vehicles from any liabilities arising from accidents involving their vehicle, contrary to the motor commercial policy which is designed for businesses and for vehicles that are being utilised for commercial purposes (Taveo, 2023). Any vehicle owner in Kenya cannot freely operate the vehicle without motor vehicle insurance because motor third party risks cover is mandatory and driving without insurance is considered illegal. Hence the three classes of motor insurance: third-party only insurance (TPO), third-party fire and theft (TPFT) and comprehensive insurance. All three include coverage for damage or injury to third parties when an accident occurs. TPFT extends the basic third party cover and incorporates protection against damages from a fire or theft of the insured's vehicle. Comprehensive insurance shields the insured against damages on their vehicle on top of the basic TPO coverage.

The motor vehicle insurance policies are the second largest businesses within the insurance companies in Kenya in relation to the net premiums earned in the year 2022 *(see figure 1)*. These policies also result in the second largest claim sizes relative to other products sold in the general insurance industry *(see figure 1)*. This is because every vehicle owner in Kenya is expected to take on insurance policies for their vehicles and each of these policies could result to more than one claim of varying amounts in each policy year. This only emphasises the need to have a clear and accurate forecast in order to set aside enough reserves for when a claim needs to be settled, and assess the need for motor reinsurance arrangements to spread out the risk.

Class	Net Earned Premium	Net Claims Incurred
Aviation	22,371	3,226
Engineering	1,063,150	679,626
Fire Domestic	1,187,180	212,419
Fire Industrial	3,326,752	1,450,843
Public Liability	1,667,041	542,573
Marine	2,168,496	738,573
Motor Private	25,877,641	19,461,157
Motor Commercial	25,105,506	18,182,290
Personal Accident	791,294	322,008
Medical	42,549,557	32,434,220
Theft	2,500,141	742,478
WIBA	7,753,713	2,211,911
Miscellaneous	1,951,238	804,576
Total	**115,964,081**	**77,789,964**

Figure 1: The class-wise distribution of net earned premiums and net claims incurred in 2022 by General Insurance Companies in Kenya. Retrieved from the AKI Annual Report 2022 (Page 49)

Actuaries currently use a number of techniques to forecast motor private claim amounts and frequency. The conventional methods of forecasting claims frequency and severity use existing and pre-defined probability distribution functions using historical data. Chain Ladder Method (CLM) is another method used to forecast the amount of reserves that an insurance company should keep in order to cover future claims by forecasting the past claims into the future (Kagan, 2023). It computes the incurred but not reported (IBNR) losses using run-off triangles, assuming that the past claims experiences are a good representation of the future outcome. Actuaries also use the Expected Loss Ratio (ELR) method to determine the projected amount of insurance claims relative to the premiums earned. This technique is mainly used when there is no appropriate past claims experience information for the insurer to use. Two actuaries, Bornhuetter and Ferguson, in 1972, developed another widely used method that is used to estimate an insurance company's losses, which are the claims paid out in this context. The Bornhuetter-Ferguson method is used to estimate the incurred but not reported losses for insurers by bringing together aspects of the chain ladder method and those of the expected loss ratio methods and gives weights to the proportion of losses that have already been settled and to the losses incurred (Kagan, 2023).

In as much as the above methods may be simple, straightforward and easy to use, they are not flexible and may be prone to distortion if at all the claims reporting pattern changes. The do not fully capture the volatile nature, and unpredictability of insurance claims. Mills (2009) explains that actuaries only use a few of the forecasting methods that are largely used by the general business community, which are thought to be potentially relevant for actuarial work, most of which consists of the time series forecasting techniques such ARIMA, SARIMA, exponential smoothing, and econometric modelling, among others.

Data collected over time for motor private claim experiences in Kenya from the year 2013 to 2022 fits the definition of time series data; they are collected and documented regularly over time by the IRA, and exhibit trends, seasonality, and irregularity, hence the assumption that an accurate claims prediction can be generated using time series analysis techniques.

1.2. Problem statement

The key risk for insurers within the motor vehicle insurance space in Kenya is that there is a high up-take of motor insurance policies, and that each individual policy could lead to more than one claim of an unknown amount in each policy year. This means that insurance companies are expected to be in a position to comfortably pay out the varying claim amounts reported by any of the many policyholders. This requires an accurate and prudent motor private insurance claims experience and their severity in order to charge reasonable premiums and to allow for adequate reserving. Actuaries, in attempts do this, ignore the forecasting techniques that have been used successfully with great results by general businesses because they are deemed as either too simplistic, or too complicated, but could potentially be of relevance when applied to actuarial work for ease, accuracy and flexibility. Actuaries often opt to fit historical data into existing probability distribution models, use simple moving average models, the chain-ladder method, Bornhuetter-Ferguson method, loss ratios and Bayesian forecasting models, among others, and largely ignores the time series analysis and forecasting methods. The education for actuaries and the IRA claims reserving requirements in Kenya (IRA Act Cap. 487/16) are also too fixated on the methods mentioned above, and mentions little or even nothing of other methods that could potentially yield reliable results, hence tying actuaries to specific methods of forecasting only (Dahms, 2022).

A forecast of motor private insurance claims strongly depends on past data and its rich characteristics. The ARIMA model offer flexibility of capturing various patterns and behaviours in data, making it well suited for the job that actuaries do when attempting to use past data to forecast future possibilities. This study checks the viability of the one of the time series analysis and forecasting methods, the ARIMA model, in forecasting of motor private insurance claims experiences since claims data, just like any other data collected sequentially over time, qualifies as time series data, exhibiting seasonality, trends, and irregularity, which can intuitively be deemed as viable for forecasting using time series techniques.

1.3. Objectives of the study

1.3.1. General objective

The general objective of the study is to analyse and give a five year forecast (2023 to 2027) of the motor private insurance claim amounts using the ARIMA model with data spanning from the year 2013 to 2022.

1.3.2. Specific Objectives

i. To analyse the trend and general behaviour of the motor vehicle insurance claim amounts for the top five general insurance companies in Kenya by market share as of 2022.
ii. To determine the optimal order of the parameters p, d and q for the ARIMA model that best fits the motor private insurance claims experience from 2013 to 2022 for each insurance company.
iii. To forecast the motor private insurance claim amounts for the years 2023 to 2027 using the ARIMA models for each of the five companies.
iv. To assess the viability and accuracy of using the generated optimal ARIMA model for forecasting motor private insurance claims.

1.4. Significance of the study

We have already established that general insurance companies need an accurate and prudent forecast of their possibly future claims in order to establish how much they have to reserve, how much premiums they have to charge, and if they would like to plan for reinsurance, among other important decisions, just so that they can ensure that they are solvent and that they are making profits (Poufinuas et. al, 2023).

Very little has been researched on the use of ARIMA models, or any other time series forecasting approach to make forecasts within the motor private insurance space. Having this research replicated in Kenya will provide insights tailored to the Kenyan market in terms of the accuracy of the forecasts and the viability of the use of time series analysis approaches to forecast insurance claims in the Kenyan market. This ensures that the characteristics of the Kenyan market are factored in rather than making an inference on the Kenyan motor private insurance industry based on data from other markets with characteristics and economic factors that might not mirror the those in Kenya.

This study attempts to forecast motor private insurance claims experience in an industry that is fixated on using specific forecasting methods, the chain ladder, the ELR and the BF models, based on historical data. A successful study and reliable results will give room for the incorporation of another useful and reliable technique into the actuarial analysis and forecasting toolkit because the data that is often used by actuaries is time series in nature. This kind of data is best dealt with using time series analysis and forecasting methods rather than the less flexible actuarial methods.

1.5. Scope of the study

This study only looks at the motor private insurance claim experience for the top four largest general insurance companies in Kenya, with reference to the market share of the general insurance companies in 2022 according to IRA and AKI annual reports *(see figure 2)*, which are Old Mutual General (8.7%), APA General Insurance (8.58%), GA Insurance (8.15%), CIC General Insurance (8.11%) and Britam General Insurance (6.35%) between the years 2013 to 2022. The forecast in this study will go on for the next five years (from 2023 to 2027) using the univariate ARIMA forecasting technique.

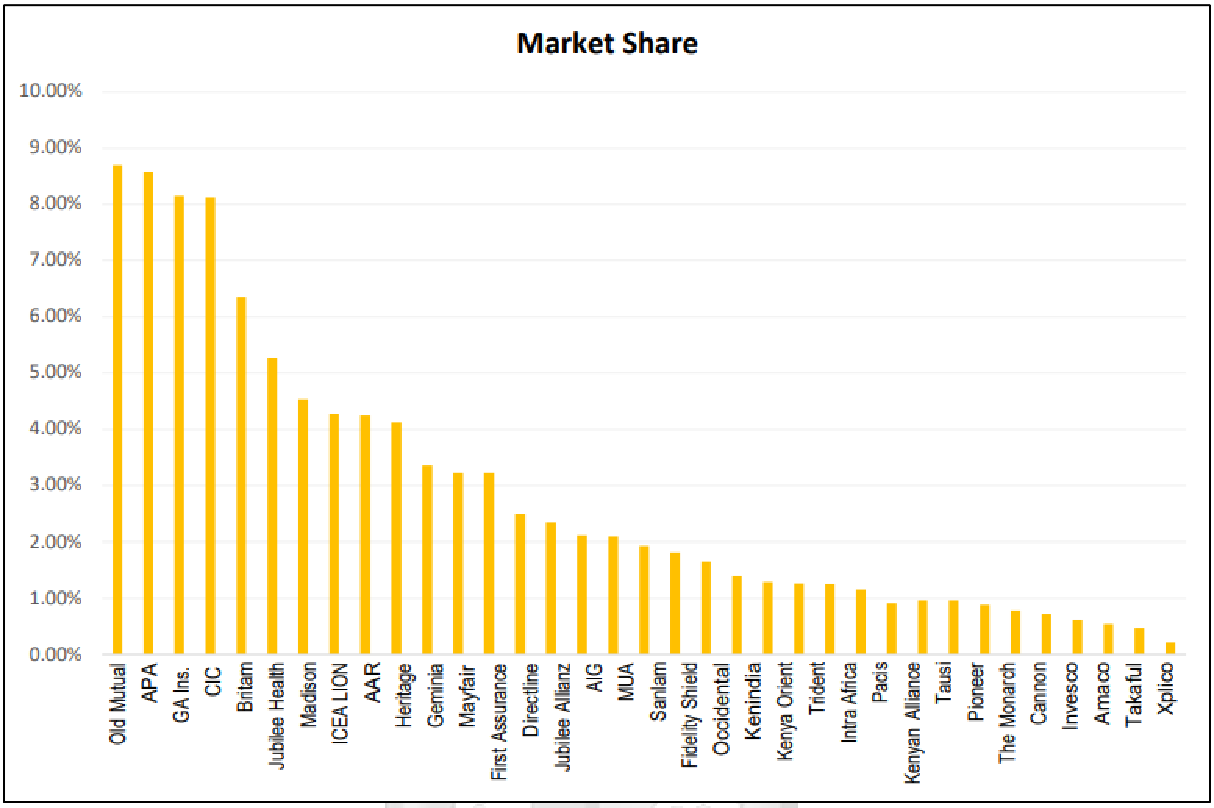


Figure 2: The market share (%) of the General Insurers in Kenya. Retrieved from the AKI Annual Report 2022 (Page 53)

CHAPTER TWO: LITERATURE REVIEW

2.1. Introduction

A proper analysis and forecast of motor private insurance claim experiences using a time series forecasting technique requires a general comprehension of the motor private claims experiences in Kenya, time series data and its components, time series analysis, and the widely recognised time series forecasting techniques and models, as well as some of the studies that other researchers have done that is similar to this study.

2.2. Motor private insurance claims experience in Kenya

Even with the insurance penetration rate in Kenya as low as about 3%, it is required by law that all motor vehicle owners have insurance for their vehicles to as protection against liabilities to third party or/and financial losses from damage or theft of their own vehicles. With this specific class of business forming the second largest portion, about 22%, of experienced claims in the general insurance companies after medical insurance claims, it leaves insurance companies with a large task of ensuring that measures are put in place to make sure that policyholders report the claims in good time, that the claims made are not fraudulent, and that there is enough reserves to settle the claims. Given the nature of motor insurance claims from the general characteristics of general insurance claims, the number of claims from each policyholder in a given policy year are not limited, and the amount of the claim is also unknown because that is often determined after the loss has occurred (Parodi, 2015). This only increases the pressure that insurers have when determining the amount of reserves to set aside, the reinsurance arrangements to opt for and the premiums they have to charge their policyholders in order to maintain solvency and profitability despite the massive claims expenses that they are subjected to, among other outflows of money necessary to put the policy in place.

The current common forecasting (reserving) methods used in the Kenyan market as prescribed and required by the IRA act are run-off (development) triangles (chain ladder, average cost per claim method, bornhuetter-ferguson method and the standard development method). They are used to estimate the future possible claims that will be reported based on those that have already been reported and/or paid. It mainly assumes that the claims loss pattern observed in the past will seemingly continue into the future. Development factors between years are computed based on the available or complete years, and the same is used it make forecasts of possible claims for the future in order to know how much to reserve for future claims settlement (Weindorfer, 2012).

2.3. The concept of time series analysis and forecasting

Time series data analysis encompasses the collection of data points at consistent intervals over a set period of time and the room to make forecasts for the future based on the trends and behaviour of the past data available. There are four major components of time series data, which inform the concept of time series decomposition. Time series decomposition into the four components help in the better comprehension of the patterns and trends in the data in order to make better forecasts. The four components are as below:

- *Trend component* (T_t) *:* This is informed by the long-run pattern or behaviour of the time series data. It may be upwards or downwards.
- *Seasonality component (* S_t) *:* This is regular, repetitive and predictable patterns at certain times of the year.
- *Cyclical component:* Are long-term wave-like patterns that occurs if the fluctuation in the data is not of a fixed or predictable frequency.
- *Irregular component (Noise)(* R_t)*:* Are random and unpredictable variations in the data that do not subscribe to a particular pattern. They are often short-term.

and the two methods of decomposition are as below based on the following assumption using the above notations:

$$\boldsymbol{y_t = f(S_t, T_t, R_t)} \qquad \textit{(1)}$$

where $\boldsymbol{y_t}$ is the data at period t;

- Additive decomposition, which is suitable when the seasonal variation is relatively constant over time.

$$\boldsymbol{y_t = S_t + T_t + R_t} \qquad \textit{(2)}$$

- Multiplicative decomposition, which is relevant when the seasonality variation increases over time.

$$\boldsymbol{y_t = S_t \times T_t \times R_t} \qquad \textit{(3)}$$

Some of the models that are commonly used to forecast time series data include the ARIMA (p, d, q) model, the exponential smoothing and the SARIMA $(p, d, q)(P, D, Q)n$ model. More of each is explained in the next chapter.

2.4. Empirical review

Similar studies have been conducted over the years to explore the application of time series models (ARIMA models) to forecast motor insurance claims experiences across different regions and contexts. These studies often employ the ARIMA model due to its robustness and flexibility in handling various types of time series data, while others integrated the ARIMA models with machine learning techniques and existing deterministic models to enhance their predictive power.

Cummins and Griepentrog (1985) tested the performance of ARIMA and econometric models alongside methods used in the insurance industry to forecast two paid claim cost series in their study within the USA using data from 1974 to 1983. The findings suggested insurers in the automobile space could potentially improve their predictions of claims arising from liabilities to third parties by the insured related to property damage by taking advantage of the existing the econometric models. For claims arising from liabilities to third party related to body injuries, the accuracy obtained from using the econometric and insurance industry models are almost similar and together, they outperform the ARIMA model, hence accurate results can be obtained by adopting the use of econometric models. This is relevant to our study because the motor private insurance can either be comprehensive, which covers damage to third party vehicle and damage to own vehicle, or third party, which takes care of damage to third party vehicle. Claims by third party in both cases qualify as property damage liability claims.

Harvey and Fernandes (1989) in their study on "Time Series Models for Insurance Claims" suggested a way to use time series models to predict future values of insurance claims using past by constructing a structural time series model that combines the claim size experience and the frequency of claims, therefore treating the aggregate claim size experience as a random sum and providing a robust framework for forecasting insurance claim predictions. The study focuses on decomposing the claim amounts into the components of time series data: seasonality, trend and the stochastic component. The model uses both the Gaussian and Poisson models, with the Poisson model accounting for the claim frequency. The model allows for the incorporation of the explanatory variables that affect the frequency and claim experiences such as: economic indicators, seasonal effects, changes in policies and regulations for accuracy of the results. One of the key limitations of this model is the availability of individual claim amounts, without which, there arises a problem with accurately estimating model parameters and validating the performances. Secondly, since the Poisson model is limited by its requirements for time varying level component while insisting on a deterministic trend and seasonality component which limits the model's flexibility in capturing complex patterns in the data. Thirdly, computational complexity arises when the Poisson model is used with large datasets or models with models that cater for too many explanatory variables. Also, dealing with zero observations and overly scattered data results to challenges with parameter estimation. These limitations highlight areas for further research and refinement.

Kumar et al. (2020), in their study, used the ARIMA model to forecast the future claim amounts of own damage motor vehicle insurance claims in India using past secondary data for a duration of 36 years (1981 to 2016). Stationarity had to be achieved, and the standard deviation, the ACF and the PACF were computed to determine the values of p, d and q. ARIMA (*1, 0, 1*) was found to be the best fit with a 70% significance with a relatively small MAPE and a relatively large, adjusted R-squared, suggesting that it would be the best for forecasting own damage motor vehicle insurance claims for the next 14 years (2017 to 2030). From the results, it was concluded that insurers would find the use of ARIMA models with the correct orders of p, d and q effective for accurate forecasts and proper future planning.

This study aims to replicate aspects of the above studies on the Kenyan motor private insurance market by performing a forecast of insurance claims using the time series ARIMA model.

CHAPTER THREE: METHODOLOGY

3.1. Introduction

This section explains the approach taken and the models used to conduct the study. It outlines the type of research carried out, the data used, the source of the data used, the models in which the data was fitted into, the data analysis software used, and the method used to validate the models used.

3.2. Data and research design

The data used is secondary data that is publicly available. The amount of incurred claims under the motor private insurance business is obtained from the IRA annual financial reports and the quarterly claims settlement reports available on the IRA website, which are normally prepared after collection of information from different insurance companies in Kenya. The five companies selected for the study are Old Mutual General Insurance, APA General Insurance, GA Insurance, CIC General Insurance and Britam General Insurance because as of the most recent annual financial report released by the IRA (for the year 2022), they are the top five companies in terms of market share percentage and earned gross premiums. The time series data used for analysis and forecasting is from the year 2013 to 2022. The two variables of interest in the study is the time component (quarterly) and its respective incurred claim amount for the 10 years from the five general insurance companies.

The data used will be compiled and cleaned in Excel from the IRA website, then exported to R programming software for data analysis and forecasting. The model used for the forecasting of the motor private claims experience is the ARIMA model. The performance of the model will be evaluated using the Ljung-Box test which gives information on the accuracy of the model. A small value of RMSE is desired.

3.3. Exponential Smoothing Model

Exponential smoothing is a time series forecasting technique characterised by the smoothing of fluctuations in the data. It was introduced by Robert Goodell in 1956 and further added on to by Charles Holt back in 1957. It uses an exponentially weighted average of past data to forecast values for the future, giving more weight to the most recent data. The main assumption in this model is that the future values are a function of the past data. The types of exponential smoothing include: the simple exponential smoothing which assumes that the data has no trend or seasonality (stationary data), the Holt's linear exponential smoothing (double exponential smoothing) which assumes a linear trend but no seasonality, and Holt-Winters' exponential smoothing (triple exponential smoothing), which assumes a trend and a seasonal component in the data (Influx Data, 2023).

A one-step ahead forecast using the exponential smoothing would use the formula below:

$$\boldsymbol{\hat{X}(1) = \alpha(X_n + (1-\alpha)X_{n-1} + (1-\alpha)^2 X_{n-2} + \dots)} \quad \textbf{\textit{(4)}}$$

The weights are α, $\alpha(1-\alpha)$, $\alpha(1-\alpha)^2$, ….

While this method is deemed simple, intuitive and highly adaptive to noise and outliers, it is limited by its simple structure and assumptions as it may not capture non-linear relationships or complex dynamics hence it might not account for structural breaks, cycles and nonlinear trends in the data. As a result, it will not be suitable for the data that is as random and complex as that of motor private insurance claims.

3.4. The Box-Jenkins ARIMA model

George Box and Gwilym Jenkins developed the model, and the concepts are documented in a 1970 publication called "Time Series Analysis: Forecasting and Control". Three key areas are considered when forecasting using the Box-Jenkins model: autoregression, differencing and moving average, denoted *p*, *d* and *q* respectively, hence is denoted *ARIMA (p , d , q).* The model was developed to provide an extensive approach to modelling time series data whilst accounting for autocorrelations in the data. Application of the model requires that the data be stationary hence the differencing component, which shows an upgrade of the ARMA *(p , q)* model. The ACF, PACF and the Akaike Information Criterion (AIC) help with the identification of the order of the autoregressive and the moving average component.

In terms of the backshift operator by Box and Jenkins (1970), the ARIMA *(p , d , q)* can be defined as:

$$\boldsymbol{(1 - \alpha_1 B - \dots - \alpha_P B^p)(1 - B)^d (X - M) = (1 + \beta_1 B + \dots + \beta_q B^q)\, e} \quad \textbf{\textit{(5)}}$$

There are three main steps to consider when using the Box-Jenkins approach to forecast: the identification of the model from the ARIMA class (the order of the ARIMA model) after making the time series stationary, the estimation of the parameter using the identified model and the diagnostic check to assess the accuracy of the model identified in the first step using the Ljung-Box 'portmanteau' test (stated below), among other techniques. The main assumption of this model is that the data is stationary before forecasts are performed i.e. the mean and variance of the data does not vary with time.

Seasonal Autoregressive Integrated Moving Average (SARIMA) model is an extension of the ARIMA model that considers the seasonality part of the model in addition to the non-seasonal part, hence is denoted SARIMA *(p , d , q) (P , D , Q) m* with *P* representing the *seasonal autoregressive, D* representing the *seasonal integrated, Q* representing the *seasonal moving average* and *m* representing the *number of observations per year.* It was developed by George Box and Gwilym Jenkins in the 1970s to address seasonality in time series which the basic ARIMA model does not consider. Appropriate application of the SARIMA model would assume that the data is stationary, and the seasonality patterns are consistent over time.

Our study will apply the ARIMA model for analysis and forecasting of the motor private insurance claims because of its flexibility in handling a wide range of data that contains very few variables. The ARIMA model has the capacity to account for the various patterns, trends, the varying volatility and any identifiable seasonal fluctuations present in the motor private insurance claims data.

3.5. Building the ARIMA model

The motor private claims experience data is plotted to check for stationarity status and for the observation and analysis of the trend, seasonality, irregularities and other observable properties in the data plot. The test for stationarity is done using the Augmented Dickey-Fuller (ADF) test with the following hypothesis:

> *H_0 : There exists a unit root in the time series hence non- stationary*
>
> *H_A : There exists no unit root in the time series hence stationary*
>
> (We reject the null hypothesis is the p-value is less than 5%.)

Without enough evidence to reject the null hypothesis, the data is then differenced to achieve stationarity, and the ACF and PACF are then used to determine the order of the ARIMA model. The test for the accuracy of the model is then performed before the forecast is generated using the Ljung-Box test (equation 6). Low values from the Ljung-Box test are desired.

$$n\,(n+2)\sum_{k=1}^{m}\frac{r_k^2}{n-k} \sim \mathcal{X}_m^2 \qquad (6)$$

CHAPTER FOUR: RESULTS AND ANALYSIS

4.1. Introduction

In this section, the results from the analysis and forecast of motor private insurance claim amount data from UAP, APA, GA, CIC and Britam General using ARIMA models is displayed and explained.

4.2.Descriptive statistics

UAP demonstrates moderate average performance with a mean of 230,833,792 and relatively low variability compared to other companies, as shown by its standard deviation of 76,404,286. However, its distribution is highly right-skewed (skewness = 2), indicating a concentration of smaller values with a few very high outliers. This is further supported by a high kurtosis value of 8, suggesting the presence of extreme observations. The range of its data is 408,124,000, with a minimum of 144,787,000 and a maximum of 552,911,000. These metrics imply that while UAP's data is somewhat consistent, it includes significant outliers that greatly influence the overall distribution. Its confidence interval at 95% is ±24,435,276, indicating a relatively narrow range of uncertainty around the mean.

APA displays balanced and symmetrical data, with a mean of 308,207,408 and a standard deviation of 77,632,165, suggesting moderate variability. The skewness and kurtosis values are both 0, indicating a perfectly symmetrical distribution with no significant outliers. The range of 339,872,000, spanning from a minimum of 102,986,000 to a maximum of 442,858,000, is consistent with its moderate standard deviation. The confidence interval at 95% is ±24,827,970, reflecting precision in the estimation of its mean. APA stands out for its lack of skewness and outliers, making its dataset relatively stable and predictable.

GA shows an average performance with a mean of 125,996,308.3 and a standard deviation of 83,037,306.23, indicating moderate variability. The distribution has slight right skewness (skewness = 0.75), suggesting a few higher values pulling the mean upwards. The kurtosis value of -0.53 indicates a flatter distribution with fewer extreme outliers compared to other companies. The range is relatively smaller at 287,132,000, with a minimum of 32,785,000 and a maximum of 319,917,000. The confidence interval at 95% is ±26,556,618.65, which is slightly wider, reflecting moderate precision in the estimation of its mean.

CIC has the lowest mean (125,996,308.3) and the smallest range (287,132,000) among the companies, indicating a lower scale of values and more consistent data. Its standard deviation of 83,037,306.23 suggests moderate variability, while a skewness of 0.75 indicates slight right skewness, with a few higher values present. The kurtosis of -0.53 reflects a flatter distribution, with fewer extreme outliers. CIC's confidence interval at 95% is ±26,556,618.65, reflecting moderate precision around the mean. Overall, CIC's data is smaller in scale, relatively stable, and less affected by outliers than some of the other companies.

Britam shows the highest average value (406,434,633.3) and the largest variability, with a standard deviation of 106,386,692.5 and a range of 577,424,000. Its data is moderately right-skewed (skewness = 1.08), suggesting a concentration of smaller values with a few high outliers, while a kurtosis value of 3.3 points to a distribution with significant extreme values. The minimum value is 215,581,000, and the maximum is 793,005,000, highlighting its large scale. The confidence interval at 95% is ±32,941,553.5, the widest among the companies, reflecting greater uncertainty in the mean estimation. BRITAM's dataset is volatile, with high variability and influential outliers.

Table 1: Summary Statistics for UAP, APA,GA, CIC and Britam

UAP		APA		GA	
Mean	230,833,791.68	Mean	308,207,408.33	Mean	125,996,308.33
Standard Error	12,080,578.34	Standard Error	12,274,722.98	Standard Error	13,129,350.92
Median	223,123,500.00	Median	314,209,500.00	Median	112,685,500.00
Mode	241,158,000.00	Mode	338,029,500.00	Mode	192,483,500.00
Standard Deviation	76,404,286.01	Standard Deviation	77,632,164.55	Standard Deviation	83,037,306.23
Kurtosis	7.71	Kurtosis	0.27	Kurtosis	- 0.53
Skewness	2.33	Skewness	- 0.48	Skewness	0.75
Range	408,124,000.00	Range	339,872,000.00	Range	287,132,000.00
Minimum	144,787,000.00	Minimum	102,986,000.00	Minimum	32,785,000.00
Maximum	552,911,000.00	Maximum	442,858,000.00	Maximum	319,917,000.00
Sum	9,233,351,667.00	Sum	12,328,296,333.00	Sum	5,039,852,333.00
Count	40.00	Count	40.00	Count	40.00
Confidence Level(95%)	24,435,275.89	Confidence Level(95%)	24,827,970.49	Confidence Level(95%)	26,556,618.65

CIC		BRITAM	
Mean	406,434,633.33	Mean	272,412,233.33
Standard Error	16,821,213.05	Standard Error	16,286,004.69
Median	393,459,000.00	Median	266,512,500.00
Mode	438,596,500.00	Mode	397,798,500.00
Standard Deviation	106,386,692.52	Standard Deviation	103,001,737.59
Kurtosis	3.30	Kurtosis	0.63
Skewness	1.08	Skewness	0.30
Range	577,424,000.00	Range	488,388,000.00
Minimum	215,581,000.00	Minimum	53,218,000.00
Maximum	793,005,000.00	Maximum	541,606,000.00
Sum	16,257,385,333.00	Sum	10,896,489,333.00
Count	40.00	Count	40.00
Confidence Level(95%)	34,024,114.59	Confidence Level(95%)	32,941,553.50

4.3. Analysis and forecasting results for UAP general insurance company

Figure 3 shows an initial plot of the incremental motor private incurred claims amount quarterly data for UAP from the year 2013 to 2022. There is no visible evidence of a trend upwards or downwards. There is also no obvious recurring seasonality. However, there is significant fluctuations, with sharp peaks in 2014 and around 2018. After these, the fluctuations become relatively moderate.

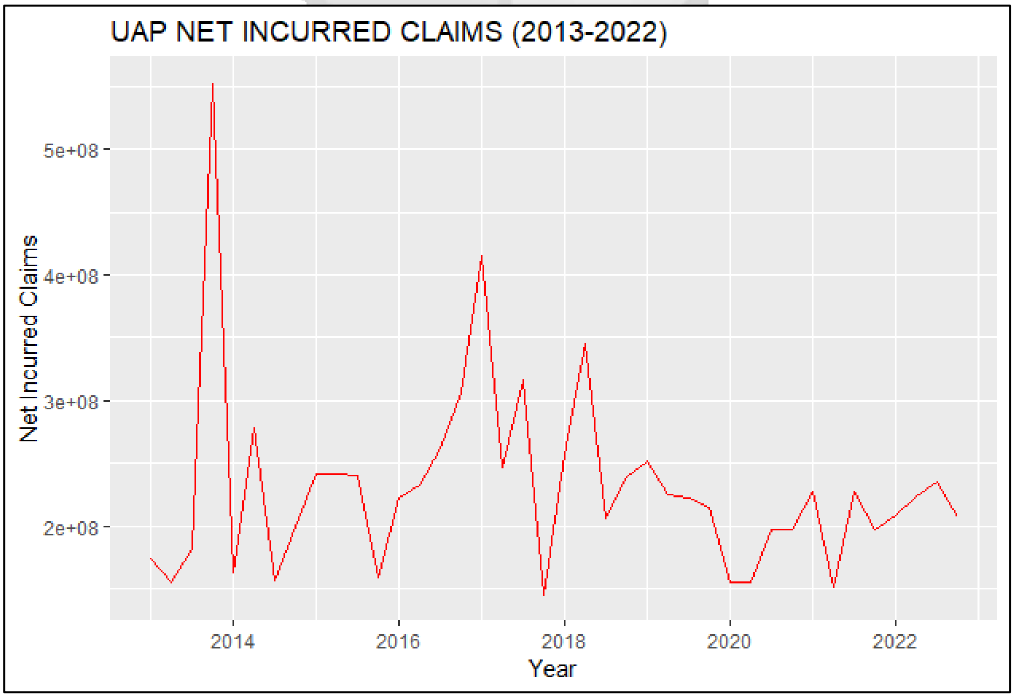


Figure 3: Initial plot for UAP incurred claims data

The series does not appear stationary because of the variability in magnitude and lack of consistency in the mean. The test for stationarity using the ADF test, with a null hypothesis stating that the series is not stationary, and the alternative stating that the series is stationary, confirmed that the data was not stationary with a p-value of 0.1208, which is greater than 0.05. Consequently, the first differencing (see figure 4) was done, giving rise to a stationary dataset with p-value of 0.01 from the ADF test.

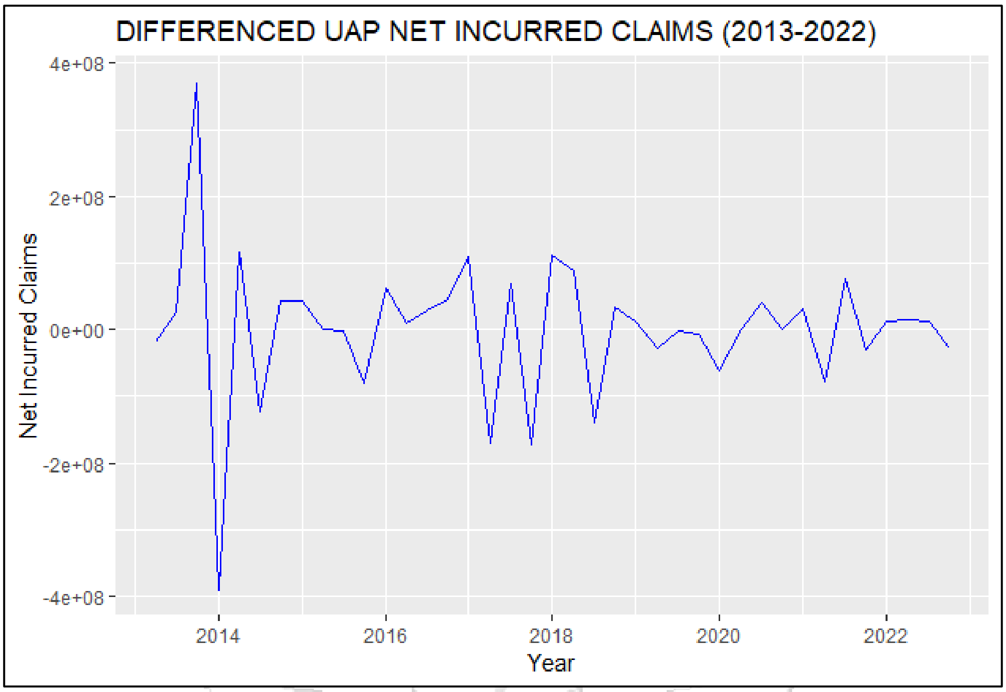


Figure 4:Plot for UAP after first differencing

The resultant ACF and PACF plots (see appendix 1) from the differences data was used to generate optimal ARIMA models for performing the forecast. Table 1 shows a list of the suggested models.

Table 2: Summary of suitable models for UAP

Parameter	**ARIMA (1,1,1)**	**ARIMA (1,1,3)**	**ARIMA (2,1,2)**	**ARIMA (1,1,2)**
Ar1	-0.0550	-0.3004	0.7015	-0.6703
Ar2	-	-	0.0890	-
Ma1	-0.9308	-0.6932	-1.7179	-0.3411
Ma2	-	-0.1326	0.7179	-0.5476
Ma3	-	-0.1735	-	-

Log-Likelihood	-765.01	-764.58	-764.87	-764.83
AIC	1536.02	1539.16	1539.74	1537.67
BIC	1541.01	1547.48	1548.06	1544.32
Ljung-Box test p-value	0.7954	-	-	-

ARIMA (1,1,1) had the lowest AIC value, suggesting that it was the optimal model of all the four suggested models. Hence ARIMA (1,1,1) was used to perform the forecast (see figure 5). The central line represents the point forecast, showing that the expected value of net incurred claims over the next 5 years (2023–2027) appears to stabilize around the average of the historical data. This suggests that no significant trend or drastic changes are expected. The shaded regions represent the prediction intervals. These intervals widen over time, reflecting greater uncertainty in forecasts further into the future. In 2023, the claims are likely to fall closer to the historical average. By 2027, the range of possible values is much broader due to compounding uncertainty.

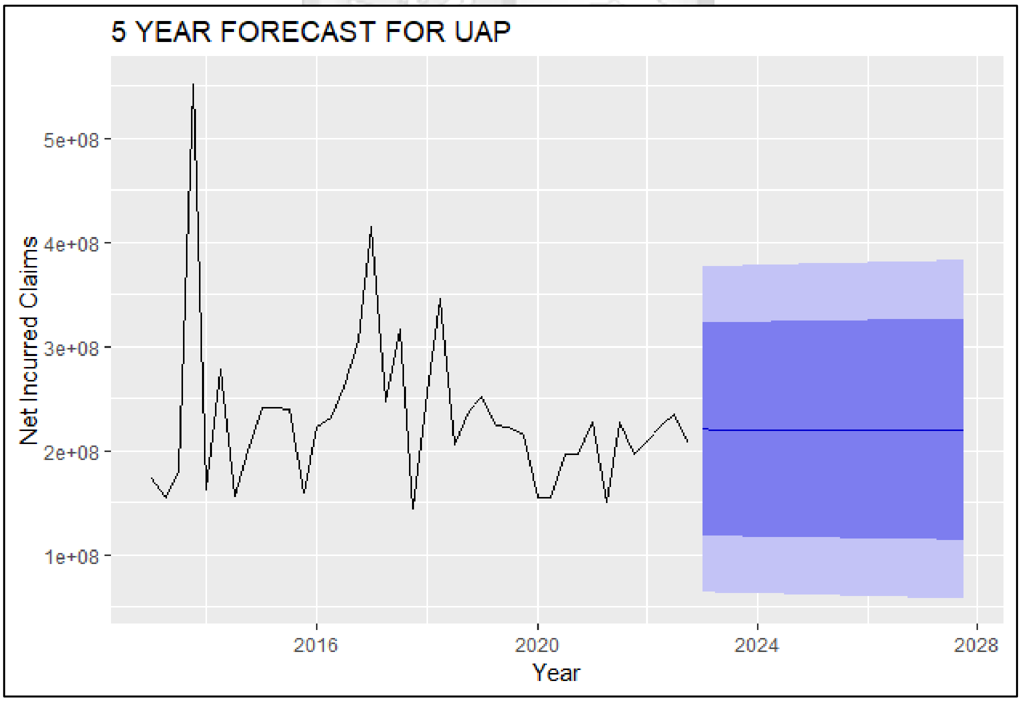


Figure 5: A 5 year forecast using ARIMA (1,1,1) for UAP

To check the model fit, the Ljung-Box test was performed, with the null hypothesis stating that the there is no significant autocorrelation at any specified lag hence the model is a good fit. The generated p-value was 0.7954, which is greater than 0.05. This suggests that ARIMA (1,1,1) is a good model for forecasting the incurred motor private claims for UAP.

4.4. Analysis and forecasting results for APA general insurance company

Figure 6 shows an initial plot of the incremental motor private incurred claims quarterly data for APA from the year 2013 to 2022. There appears to be a slight upward trend overall, suggesting a general increase in net incurred claims over the period 2013-2022. There is absence of seasonality, but irregular fluctuations are evident throughout the series.

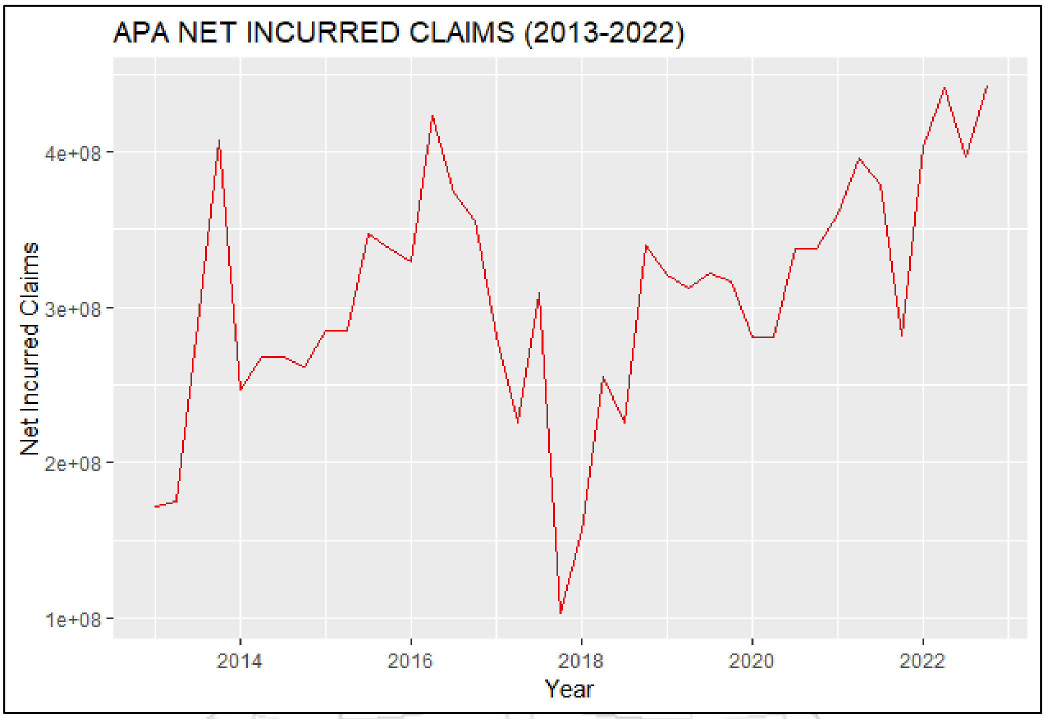


Figure 6: Initial plot for APA incurred claims data

The series does not appear stationary dure to the upward trend and the variability. A test for stationarity using the ADF test, with a null hypothesis stating that the series is not stationary, and the alternative stating that the series is stationary, confirmed that the data was not stationary with a p-value of 0.4168, which is greater than 0.05. Consequently, the first differencing was done, giving rise to a stationary dataset with p-value of 0.04 from the ADF test (see figure 7).

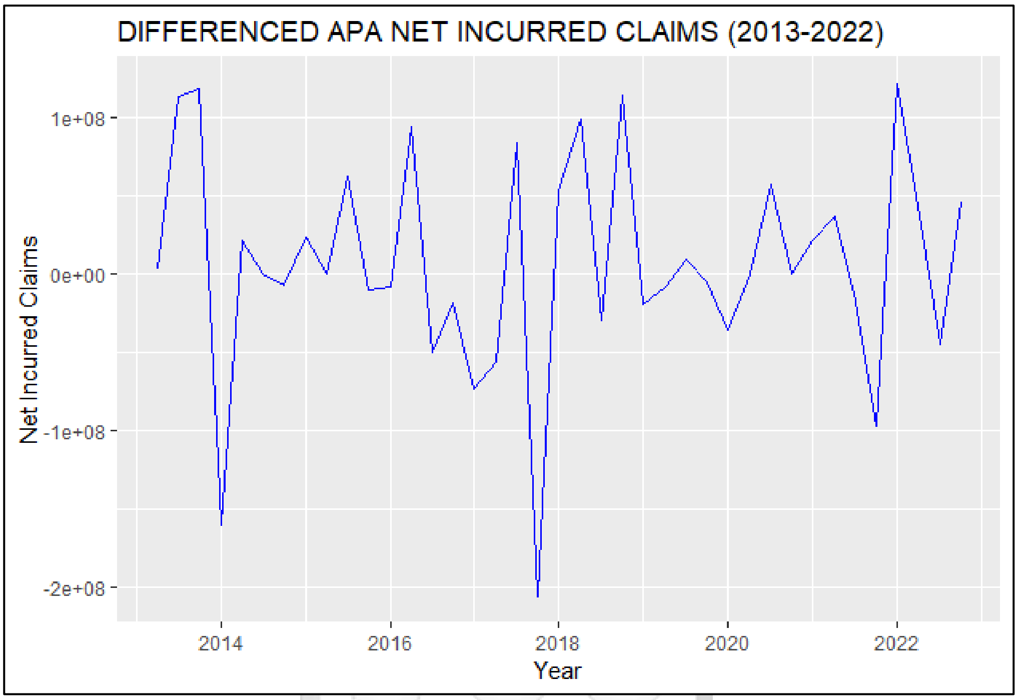


Figure 7: Plot for APA after first differencing

The resultant ACF and PACF (see appendix 2) plots from the differences data was used to generate optimal ARIMA models for performing the forecast. Table 2 shows a list of the suggested models.

Table 3: Summary of suitable models for APA

Parameter	ARIMA (6,1,0)	ARIMA (1,1,1)	ARIMA (0,1,1)	ARIMA (1,1,2)
Ar1	**-0.2303**	**0.4779**	**-**	**-0.5926**
Ar2	**-0.2839**	**-**	**-**	**-**
Ar3	**0.1451**	**-**	**-**	
Ar4	**-0.2951**	**-**	**-**	
Ar5	**-0.0696**	**-**	**-**	
Ar6	**-0.3971**	**-**	**-**	
Ma1	**-**	**-0.8417**	**-0.3843**	**0.4629**
Ma2	**-**	**-**	**-**	**-0.5371**
Ma3	**-**	**-**	**-**	**-**
Log-Likelihood	**-753.19**	**-757.1**	**-757.59**	**-754.62**
AIC	**1520.38**	**1520.2**	**1519.18**	**1517.24**
BIC	**1532.02**	**1525.19**	**1522.51**	**1523.9**

Ljung-Box test p-value	-	-	-	0.9239

ARIMA (1,1,2) had the lowest AIC value, suggesting that it was the optimal model of all the four suggested models. Hence ARIMA (1,1,2) was used to perform the forecast (see figure 8). The central line represents the forecasted Net Incurred Claims for the next 5 years. The forecast suggests that the future claims will stabilize around the mean of the stationary data, with minor variations. The shaded regions around the forecast line represent the prediction intervals. The darker area corresponds to the 80% confidence interval while the lighter area corresponds to the 95% confidence interval. These intervals widen as the forecast moves further into the future, reflecting increased uncertainty.

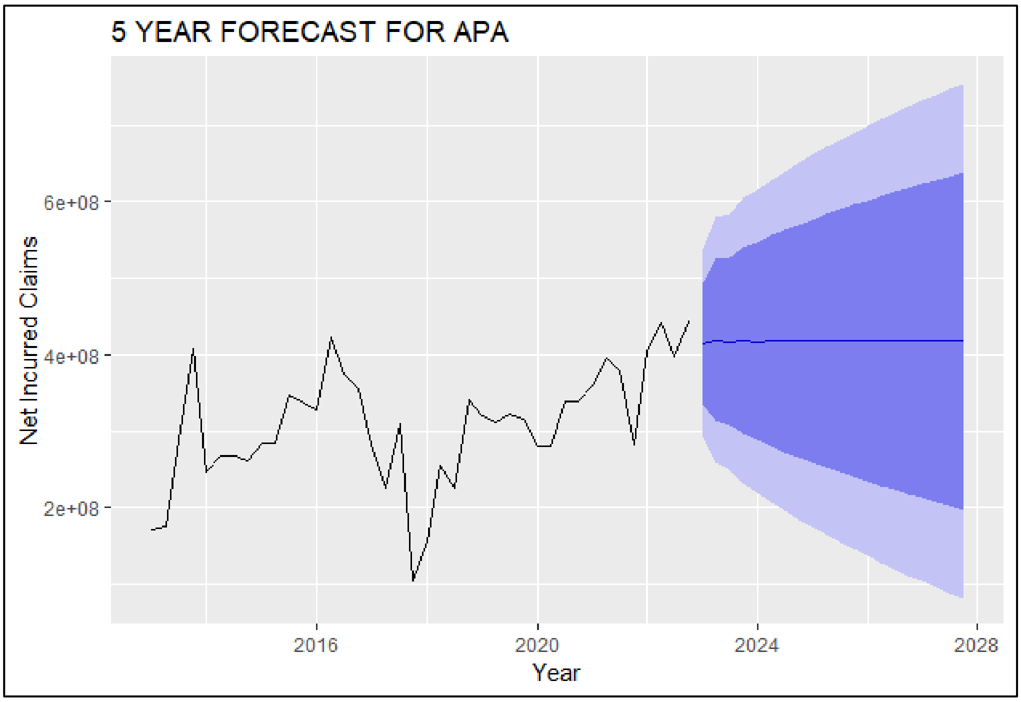


Figure 8: A 5 year forecast using ARIMA (1,1,2) for APA

To check the model fit, the Ljung-Box test was performed, with the null hypothesis stating that the there is no significant autocorrelation at any specified lag hence the model is a good fit. The generated p-value was 0.9239, which is greater than 0.05. This suggests that ARIMA (1,1,2) is a good model for forecasting the incurred motor private claims for APA.

4.5. Analysis and forecasting results for GA insurance company

Figure 9 shows an initial plot of the incremental motor private incurred claims quarterly data for GA from the year 2013 to 2022. There is a clear upward trend in the net incurred claims from 2013 to 2022, indicating that the values are generally increasing over time. While not very distinct, there appear to be recurring fluctuations, especially noticeable in the peaks and troughs across the years. Irregular fluctuations are also evident throughout the series.

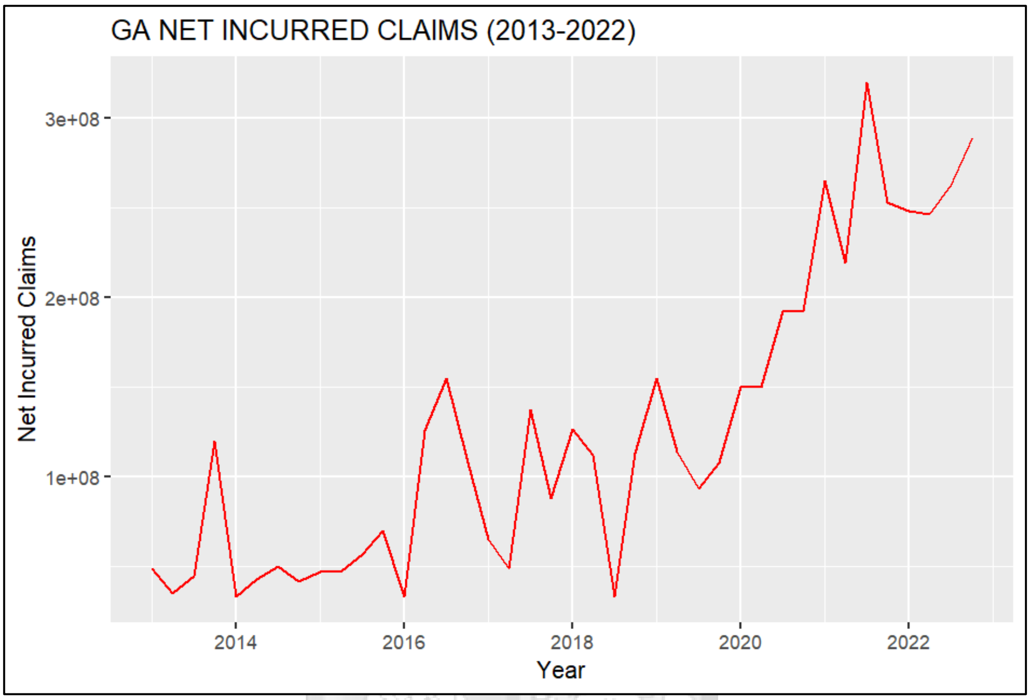


Figure 9: Initial plot for GA incurred claims data

The series does not appear stationary because of the variability in magnitude and lack of consistency in the mean. The test for stationarity using the ADF test, with a null hypothesis stating that the series is not stationary, and the alternative stating that the series is stationary, confirmed that the data was not stationary with a p-value of 0.5285, which is greater than 0.05. Consequently, the first differencing was done, giving rise to a stationary dataset with p-value of less than 0.01 from the ADF test (see figure 10).

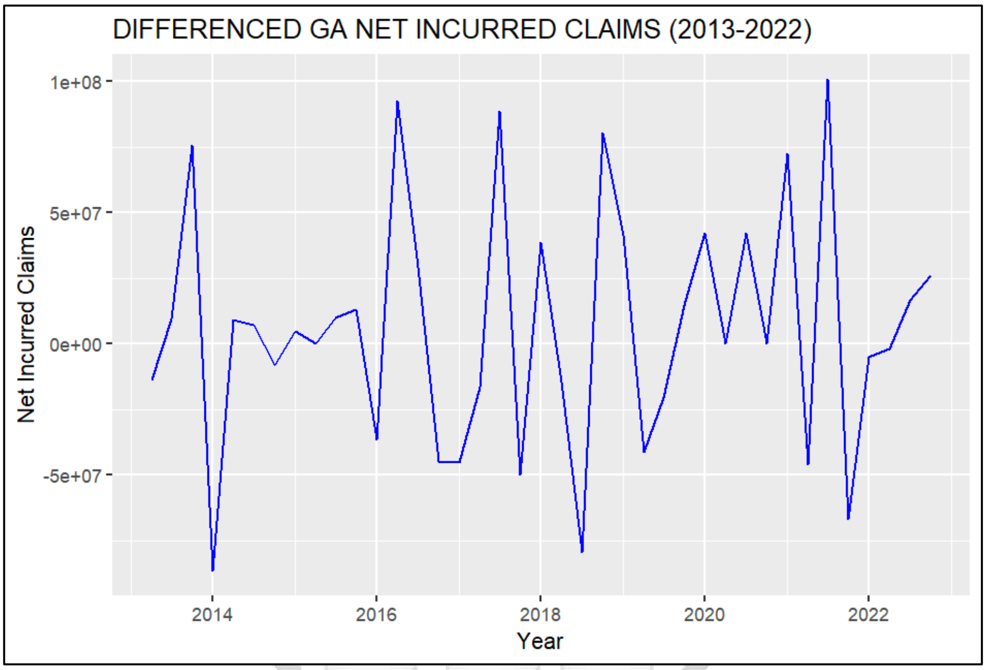


Figure 10: Plot for GA after first differencing

The resultant ACF and PACF plots (see appendix 3) from the differences data was used to generate optimal ARIMA models for performing the forecast. Table 3 shows a list of the suggested models.

Table 4: Summary of suitable models for GA

Parameter	ARIMA (1,1,1)	ARIMA (0,1,1)
Ar1	0.0117	-
Ma1	-0.4890	-0.4815
Log-Likelihood	-739.91	-739.91
AIC	1485.81	1483.82
BIC	1490.8	1487.14
Ljung-Box test p-value	-	0.8001

ARIMA (0,1,1) had the lowest AIC value, suggesting that it was the optimal model of the two suggested models. Hence ARIMA (0,1,1) was used to perform the forecast (see figure 11). The 5-year forecast for GA indicates a continued increase in Net Incurred Claims over the next five years, albeit at a slower pace than the historical trend. However, the forecast is subject to significant uncertainty, as evidenced by the widening of the forecast range towards the end of the period. This suggests that the actual claims may deviate substantially from the projected path.

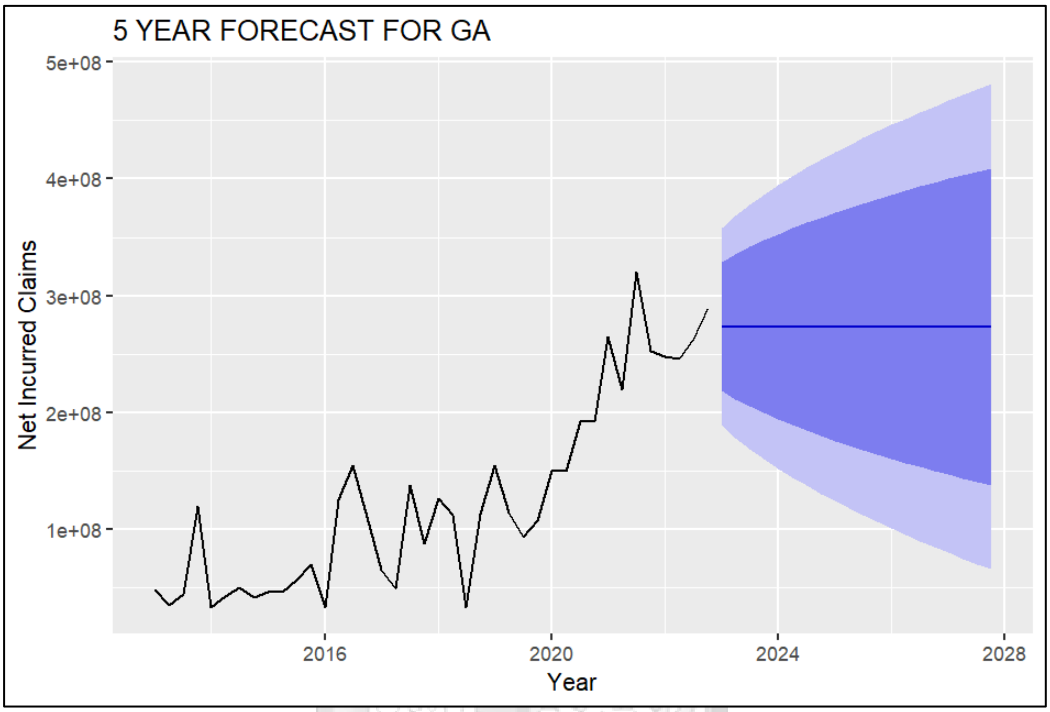


Figure 11: A 5 year forecast using ARIMA (0,1,1) for GA

To check the model fit, the Ljung-Box test was performed, with the null hypothesis stating that the there is no significant autocorrelation at any specified lag hence the model is a good fit. The generated p-value was 0.8001, which is greater than 0.05. this suggests that ARIMA (0,1,1) is a good model for forecasting the incurred motor private claims for GA.

4.6. Analysis and forecasting results for CIC general insurance company

Figure 12 shows an initial plot of the incremental motor private incurred claims quarterly data for CIC from the year 2013 to 2022. There appears to exhibit a slight upward trend over the 2013-2022 period, indicating a gradual increase in claims. While no clear seasonal pattern is discernible, there might be a hint of longer-term cycles with periods of higher and lower claims. However, a significant amount of irregularity or noise obscures the underlying trend and any potential cyclical components.

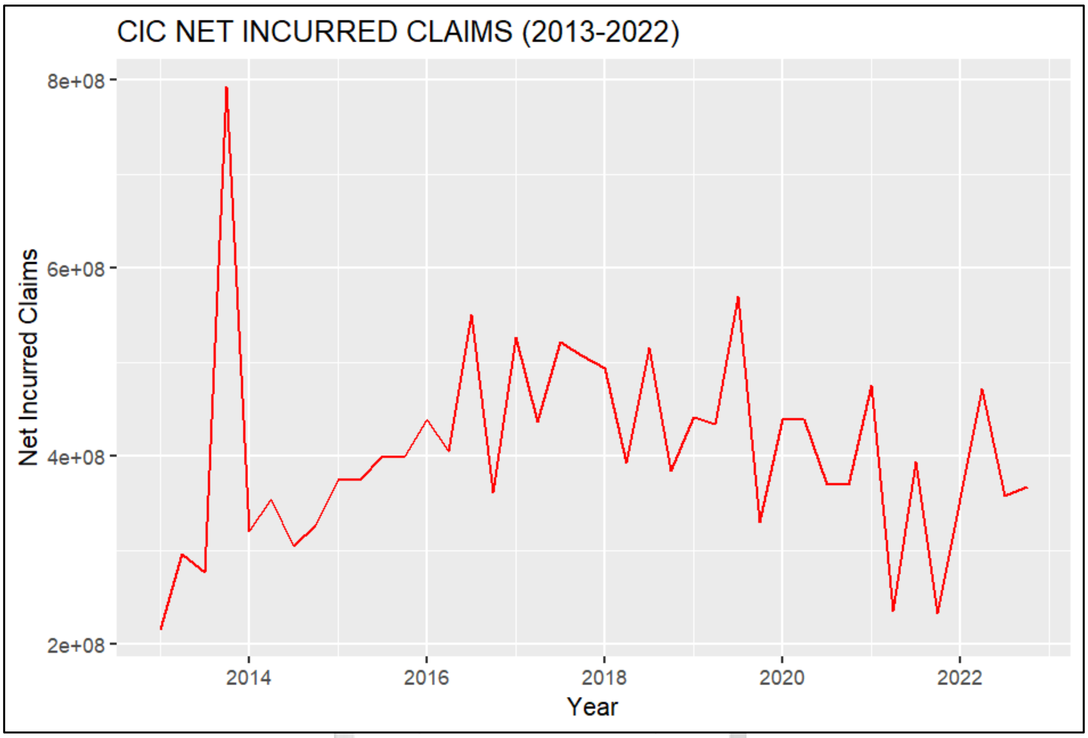


Figure 12: Initial plot for CIC incurred claims data

The series does not appear stationary because of the variability in magnitude and lack of consistency in the mean. The test for stationarity using the ADF test, with a null hypothesis stating that the series is not stationary, and the alternative stating that the series is stationary, confirmed that the data was not stationary with a p-value of 0.5803, which is greater than 0.05. Consequently, the first differencing was done, giving rise to a stationary dataset with p-value of less than 0.01 from the ADF test (see figure 13).

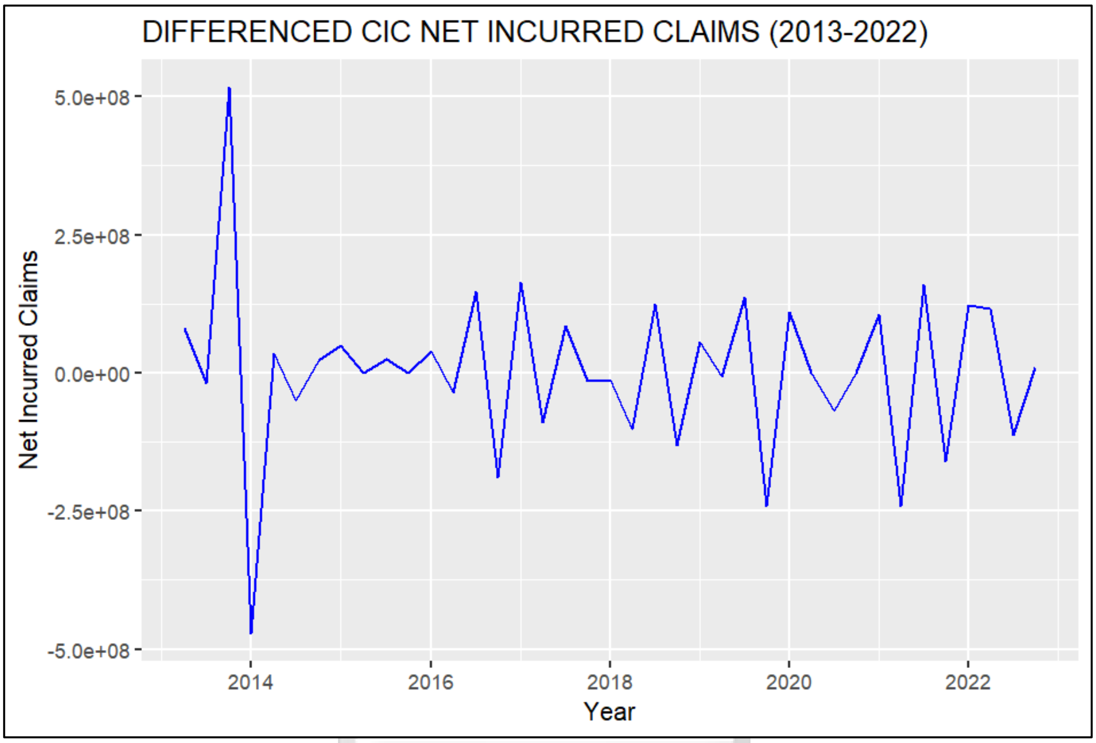


Figure 13: Plot for CIC after first differencing

The resultant ACF and PACF plots (see appendix 4) from the differences data was used to generate optimal ARIMA models for performing the forecast. Table 4 shows a list of the suggested models.

Table 5: Summary of suitable models for CIC

Parameter	ARIMA (1,1,1)	ARIMA (0,1,1)
Ar1	-0.2131	-0.3004
Ma1	-0.7830	-0.8596
Log-Likelihood	-777.12	-777.73
AIC	1560.23	1559.47
BIC	1565.22	1562.8
Ljung-Box test p-value	-	

ARIMA (0,1,1) had the lowest AIC value, suggesting that it was the optimal model of the two suggested models. Hence ARIMA (0,1,1) was used to perform the forecast (see figure 14). The 5-year forecast for CIC suggests a relatively stable level of Net Incurred Claims over the next five years, with the central forecast indicating little change. However, the forecast is subject to significant uncertainty, as evidenced by the wide range of possible outcomes.

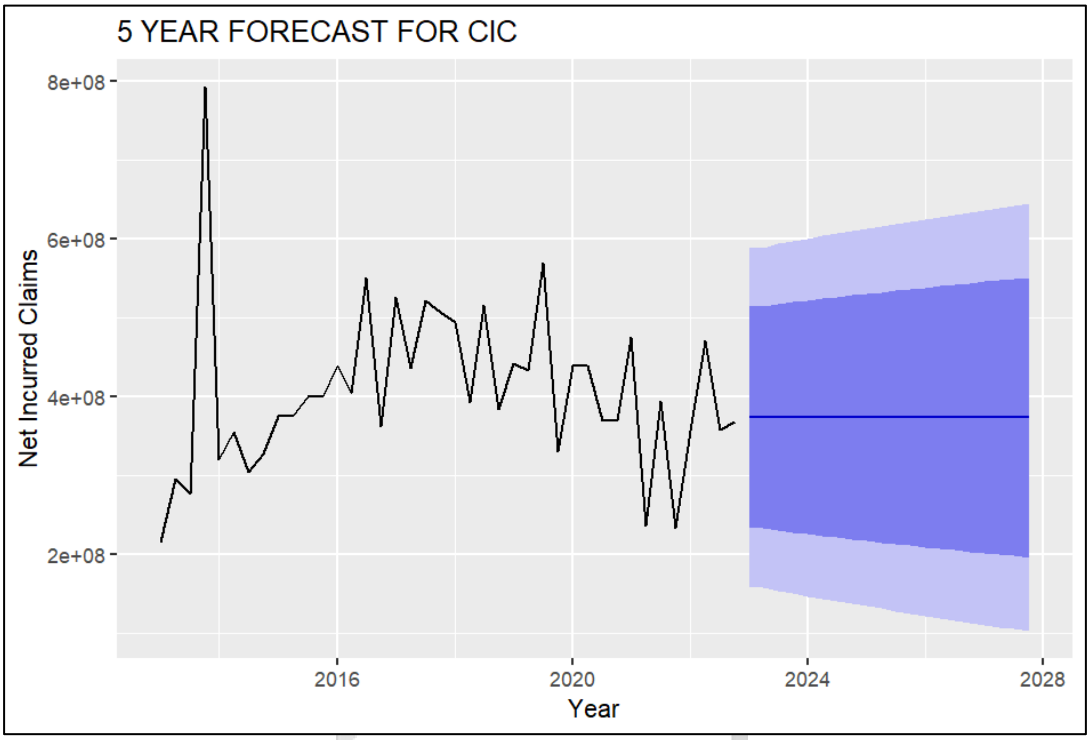


Figure 14: A 5 year forecast using ARIMA (0,1,1) for CIC

To check the model fit, the Ljung-Box test was performed, with the null hypothesis stating that the there is no significant autocorrelation at any specified lag hence the model is a good fit. The generated p-value was 0.9722, which is greater than 0.05. this suggests that ARIMA (0,1,1) is a good model for forecasting the incurred motor private claims for CIC.

4.7. Analysis and forecasting results for Britam general insurance company

Figure 15 shows an initial plot of the incremental motor private incurred claims quarterly data for Britam from the year 2013 to 2022. The graph primarily demonstrates noise. The claims data exhibits significant fluctuations without a clear discernible trend or repeating pattern. While there might be hints of short-term cycles or potential correlations with external factors, the dominant characteristic is the high level of variability and unpredictability. There is no clear evidence of a consistent trend, seasonality, or long-term cycles.

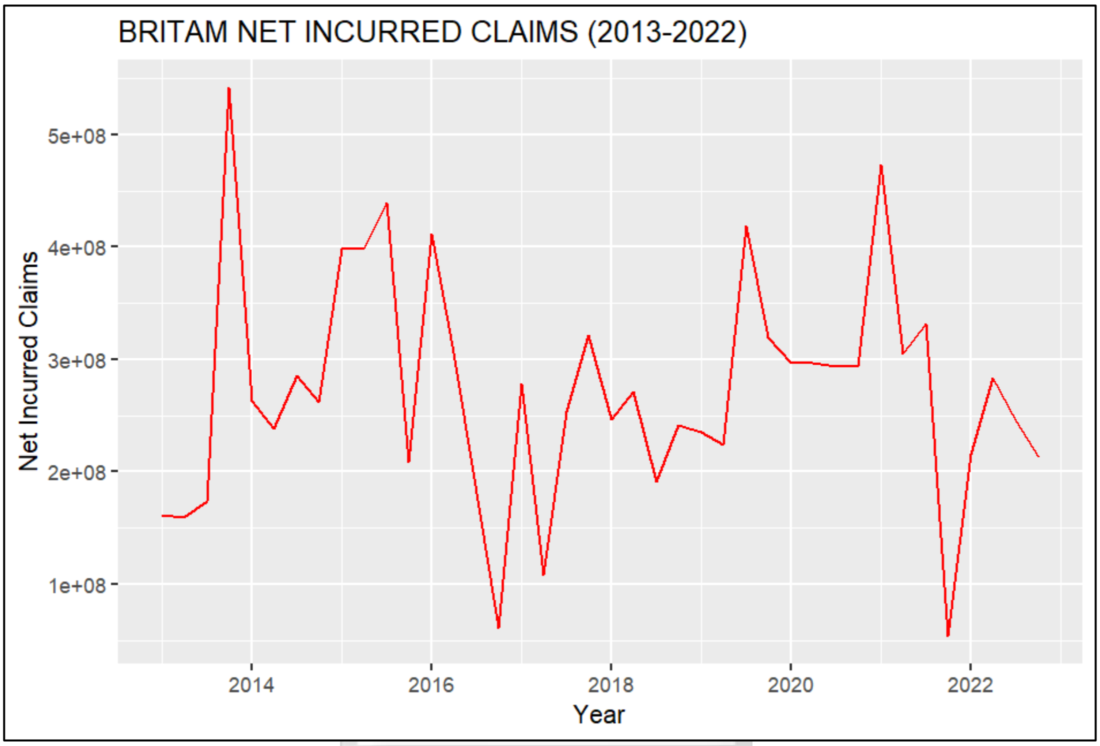


Figure 15: Initial plot for Britam Incurred claims data

The series does not appear stationary because of the variability in magnitude and lack of consistency in the mean. The test for stationarity using the ADF test, with a null hypothesis stating that the series is not stationary, and the alternative stating that the series is stationary, confirmed that the data was not stationary with a p-value of 0.3239, which is greater than 0.05. Consequently, the first differencing was done, giving rise to a stationary dataset with p-value of less than 0.01 from the ADF test (see figure 16).

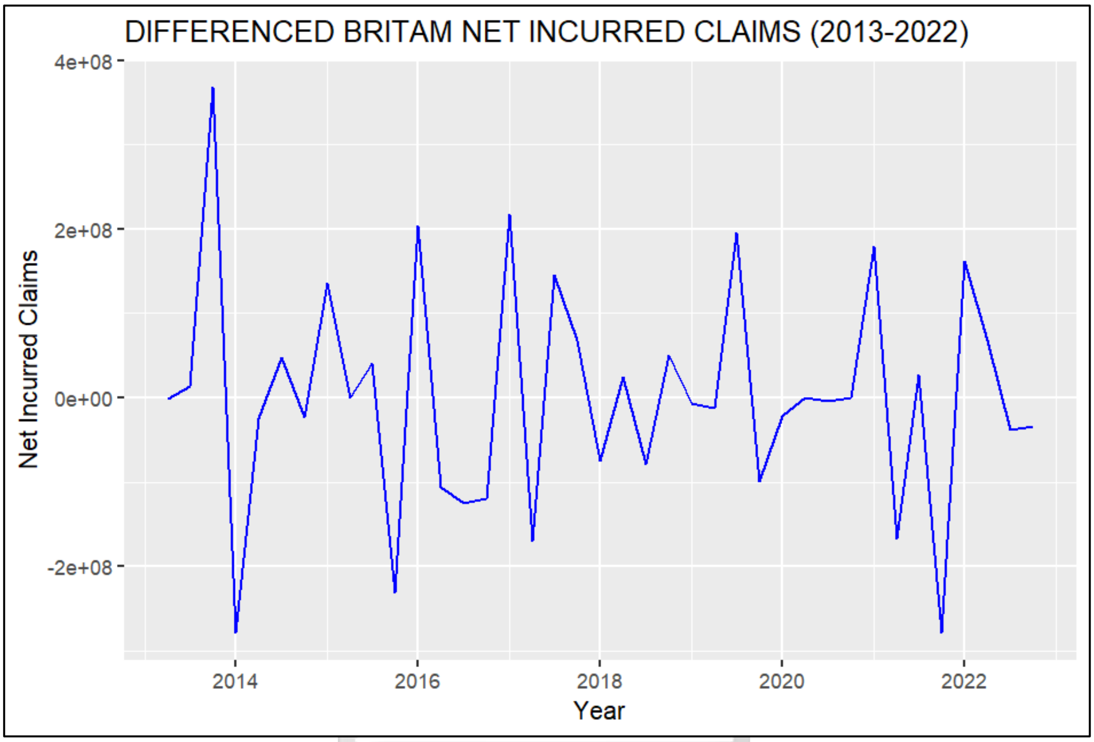


Figure 16

The resultant ACF and PACF plots (see appendix 5) from the differences data was used to generate optimal ARIMA models for performing the forecast. Table 5 shows a list of the suggested models.

Table 6: Summary of suitable models for UAP

Parameter	ARIMA (1,1,1)	ARIMA (0,1,1)
Ar1	0.1524	-
Ma1	-1.0000	-1.0000
Log-Likelihood	-776.31	-776.74
AIC	1558.61	1557.49
BIC	1563.6	1560.81
Ljung-Box test p-value	-	

ARIMA (0,1,1) had the lowest AIC value, suggesting that it was the optimal model of all the four suggested models. Hence ARIMA (0,1,1) was used to perform the forecast (see figure 17). The forecast indicates a relatively stable trajectory for Britam's Net Incurred Claims over the next five years. The projected average remains largely flat, suggesting consistent claims activity. However, the forecast is accompanied by a degree of uncertainty, represented by a shaded confidence interval. This implies that the actual claims figures could deviate from the projected average, potentially experiencing slight increases in the initial years followed by a gradual decrease towards the end of the forecast period. It's important to note that this forecast relies on certain assumptions about future conditions and that unforeseen events could significantly impact the actual outcomes.

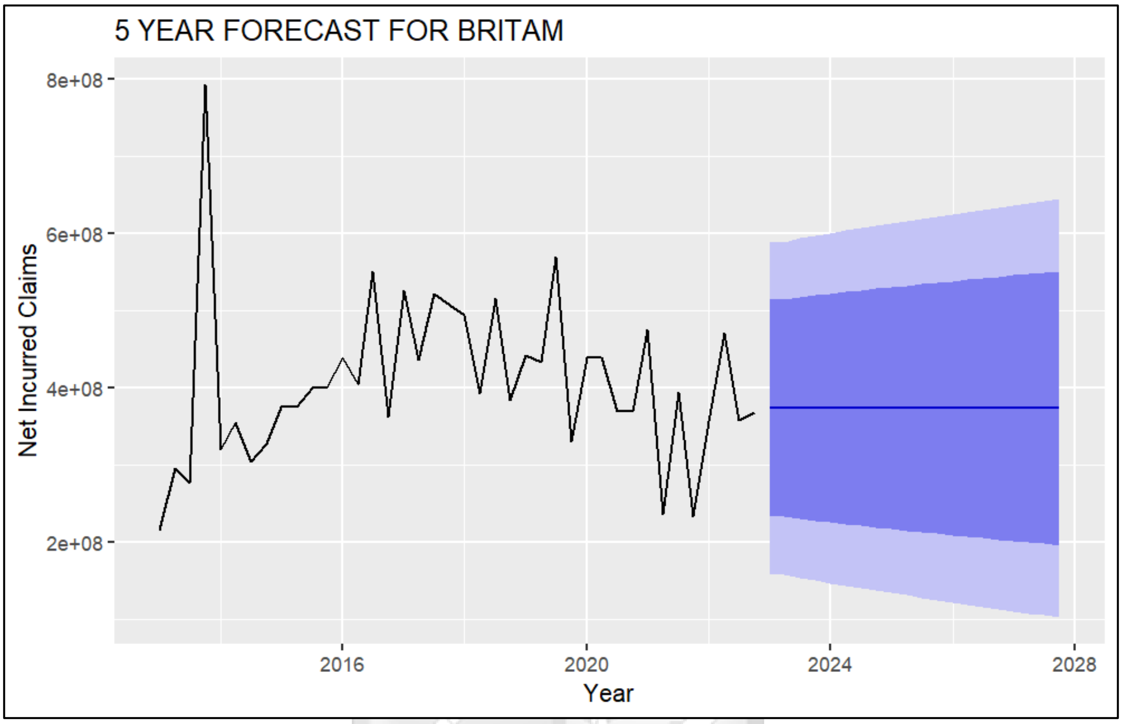


Figure 17: A 5 Year forecast using ARIMA (0,1,1) for Britam

To check the model fit, the Ljung-Box test was performed, with the null hypothesis stating that the there is no significant autocorrelation at any specified lag hence the model is a good fit. The generated p-value was 0.7954, which is greater than 0.05. this suggests that ARIMA (1,1,1) is a good model for forecasting the incurred motor private claims for Britam.

CHAPTER FIVE: DISCUSSION AND CONCLUSION

5.1. Introduction

This section discusses the results obtained in the previous chapter, and gives recommendations for insurance companies and ways in which future researchers can further improve on the study.

5.2. Conclusion

The analysis and forecasting of motor private insurance claims for UAP, APA, GA, CIC, and Britam using ARIMA models provide valuable insights into the claim patterns and expected trends for these companies. The companies exhibit varying scales and distributions of claim amounts, reflecting their operational sizes and variability. Britam has the highest mean and variability, indicating greater operational scale but also higher uncertainty and extreme observations. APA's dataset stands out for its balanced and symmetrical nature, making it relatively stable and predictable. UAP, GA, and CIC show moderate variability, with UAP and CIC influenced by right-skewed distributions and occasional extreme outliers.

All datasets were initially non-stationary, necessitating differencing to achieve stationarity for ARIMA modelling. The ADF test confirmed stationarity after the transformations. The differenced data allowed for reliable model selection and forecasting. Each company's data required specific ARIMA models for optimal forecasting. The selection criteria, including the AIC, BIC, and diagnostic checks, ensured robust model fit. UAP's claims forecast using ARIMA (1,1,1) suggests stabilization around the historical average, with increasing uncertainty over time. APA's claims, modelled with ARIMA (1,1,2), also indicate stabilization with minor variations, supported by the model's good fit based on diagnostic tests. GA's forecast using ARIMA (0,1,1) predicts a slower upward trend compared to historical increases but remains subject to significant uncertainty. CIC's claims, modelled with ARIMA (0,1,1), are expected to remain stable, with a wide prediction interval highlighting future uncertainty. Britam's forecast using ARIMA(0,1,1) suggests a relatively stable trajectory, accompanied by a certain degree of uncertainty with the future claims forecasts.

The forecasts indicate varying levels of stability and predictability across companies. The widening prediction intervals across all companies highlight increasing uncertainty in long-term forecasts, emphasizing the need for regular data updates and model refinements. These insights can guide strategic decision-making, including risk management, pricing strategies, and reserve allocations. From the analysis, it is evident that claim amounts data exhibits some of the time series components, mostly trends and irregularities, that can be accommodated for when forecasting using time series approaches, particularly the ARIMA models. The suggested optimal ARIMA models for the five general insurance companies generated a p-value of more than 5%, suggesting that the models were fit for claim amounts forecasting.

5.3. Recommendation

Insurance companies should adopt time series forecasting methods when forecasting their net incurred claims when setting aside their reserves and additional capital because the approach is straight forward and can account for the trends and noise in the data, while giving forecasts at different confidence levels to deal with the risk of future uncertainties.

Given the nature of claims and the conditions that trigger payouts by insurance companies, it is almost impossible to fully rely on past claims experience to come up with a picture of what future experiences could look like, hence the shaded regions in the forecasts to cater for the uncertainties. Recognizing this inherent unpredictability, future research efforts should delve deeper into stochastic approaches to claims forecasting. These methods, which incorporate elements of randomness and probability, can better capture the inherent uncertainty surrounding future claims events. By exploring techniques such as Monte Carlo simulations, Bayesian networks, and time series models with stochastic components, researchers can generate a wider range of potential future scenarios and assess the associated risks more comprehensively. This will ultimately lead to more robust and informative claims forecasts that better equip insurers to navigate the complexities of the future.

5.4. Limitations of the study

This study uses the univariate approach to forecast the incurred claims amounts for UAP, APA, GA, CIC and Britam. In as much as the univariate time series approach to claims forecasting generates an optimal model for each insurance company in the study, it is important to note that claims forecasting is also dependant on a number of other factors, which were not included in this study. Factors such as historical loss ratios, claim frequency and severity trends, underwriting and product mix changes, economic conditions, industry trends, catastrophe risk, legal and regulatory developments, all play a crucial role in developing accurate and robust claims forecasts for general insurance companies. Future studies should therefore incorporate these factors in order to generate more accurate and informed forecasts for claim amounts.

REFERENCES

Association of Kenya Insurers. (n.d.). *Insurance in Kenya*. Retrieved from Association of Kenya Insurers: https://www.akinsure.com/insurance-in-kenya#:~:text=General%20insurance%20is%20short%20term,commercial%20(for%20commercial%20entities).

Cummins, D. J., & Griepentrog, G. L. (1985). Forecasting automobile insurance paid claim costs using econometric and ARIMA models. *International Journal of Forecasting*, 203-2015. Retrieved from https://www.sciencedirect.com/science/article/abs/pii/0169207085900032

Dahms, R. (2022, October 7). *Why the hell do we stick to Chain Ladder and Bornhuetter-Ferguson.* Retrieved from Casualty Actuarial Society: https://www.casact.org/sites/default/files/2022-11/chain%20ladder.pdf

Express Analytics. (2021, November 9). *Everything You Need to Know About Time Series Analysis – A Primer*. Retrieved from Express Analytics: https://www.expressanalytics.com/blog/time-series-analysis/

Geek for Geek. (2023, October 20). *Time Series Decomposition Techniques*. Retrieved from Geek for Geeks: https://www.geeksforgeeks.org/time-series-decomposition-techniques/

GlobalData. (2024, March). *Kenya General Insurance Market Size and Trends by Line of Business, Distribution, Competitive Landscape and Forecast to 2028.* Retrieved from Global Data Report Store: https://www.globaldata.com/store/report/kenya-general-insurance-market-analysis/

Harvery, A., & Fernandes, C. (1989). Time Series Models for Insurance Claims. *Journal of the Institite of Actuaries*, 513-528.

Influx Data. (2023, May 24). *Exponential Smoothing: A Beginner's Guide to Getting Started*. Retrieved from Influx Data: https://www.influxdata.com/blog/exponential-smoothing-beginners-guide/#:~:text=The%20technique%20was%20first%20introduced,observations%20than%20to%20older%20observations.

Kagan, J. (2023, August 26). *Bornhuetter-Ferguson Technique: Overview, Calculations*. Retrieved from Investopedia: https://www.investopedia.com/terms/b/bornhuetterferguson-technique.asp#:~:text=The%20Bornhuetter%2DFerguson%20technique%20is%20a%20method%20of%20estimating%20incurred,to%20the%20chain%2Dladder%20method.

Kagan, J. (2023, August 9). *Chain Ladder Method (CLM) Definition, Steps to Apply It*. Retrieved from Investopedia: https://www.investopedia.com/terms/c/chain-ladder-method-clm.asp

Kenton, W. (2020, November 30). *Seasonality: What It Means in Business and Economics, Examples*. Retrieved from Investopedia: https://www.investopedia.com/terms/s/seasonality.asp#:~:text=What%20Is%20Seasonality%3F,is%20said%20to%20be%20seasonal.

Kumar, P., Kumar, S., Dipak, S., & Haragopal, V. V. (2020). Forecasting Motor Insurance Claim Amount using ARIMA models. *AIP Conference Proceedings* (pp. 1-10). AIP Publishing.

Mills, A. (2009, September). Introduction to Forecasting Methods for Actuaries. *Forecasting and Futurism*, pp. 5-9.

Mtek Services. (2024, February 19). *Understanding Motor Insurance Covers.* Retrieved from Mtek Services: https://mtek-services.com/understanding-motor-insurance-covers/

Optalitix. (2022, August 3). *Create A Claims Forecasting System*. Retrieved from Optalitix: https://www.optalitix.com/insights/use-case-claims-predictions#:~:text=The%20claims%20forecast%20is%20very,using%20the%20chain%20ladder%20method.

Parodi, P. (2015). *Pricing in General Insurance.* Florida: CRC Press: Taylor and Francis Group.

Poufinuas, T., Gogas, P., Papadimitriou, T., & Zaganidis, E. (2023, September 18). Machine Learning in Forecasting Motor Insurance Claims. *Risks*, pp. 164-183.

Ritu, S. (2023, June 8). *Introduction to SARIMA Model*. Retrieved from Medium: https://medium.com/@ritusantra/introduction-to-sarima-model-cbb885ceabe8

Ritu, S. (2023, May 12). *Tests for Stationarity in Time Series — Dickey Fuller Test & Augmented Dickey Fuller(ADF) Test*. Retrieved from Medium: https://medium.com/@ritusantra/tests-for-stationarity-in-time-series-dickey-fuller-test-augmented-dickey-fuller-adf-test-d2e92e214360

Scott, G. (2022, August 30). *Box-Jenkins Model: Definition, Uses, Timeframes, and Forecasting*. Retrieved from Investopedia: https://www.investopedia.com/terms/b/box-jenkins-model.asp#:~:text=The%20Box%2DJenkins%20Model%20forecasts,p%2C%20d%2C%20q).

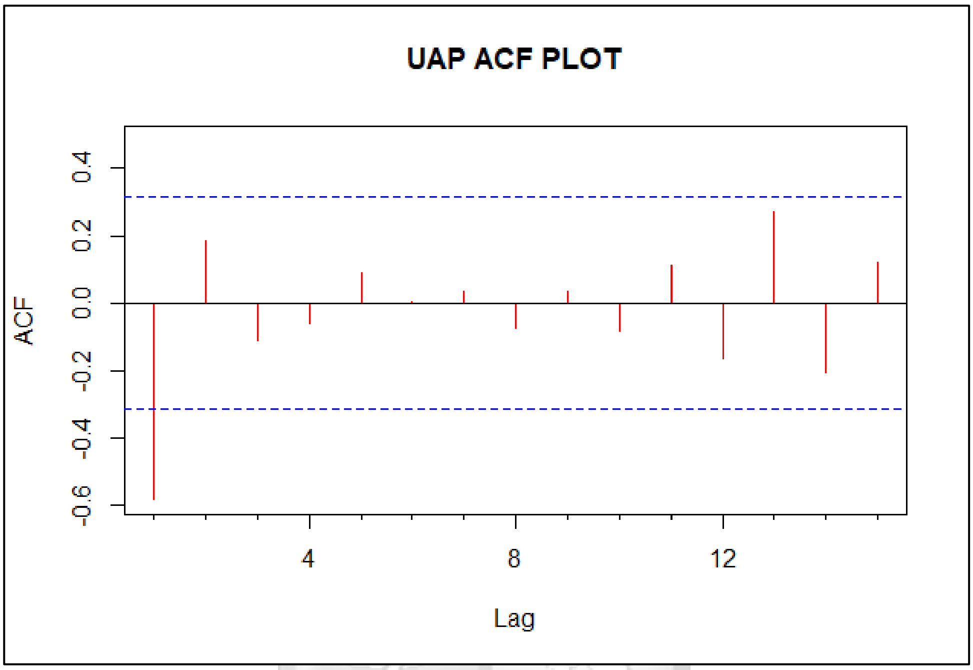
Taveo. (2023, February 3). *Motor Insurance: What Are the Differences.* Retrieved from Taveo: https://taveo.com/resources/blog/commercial-personal-motor-insurance-the-differences/

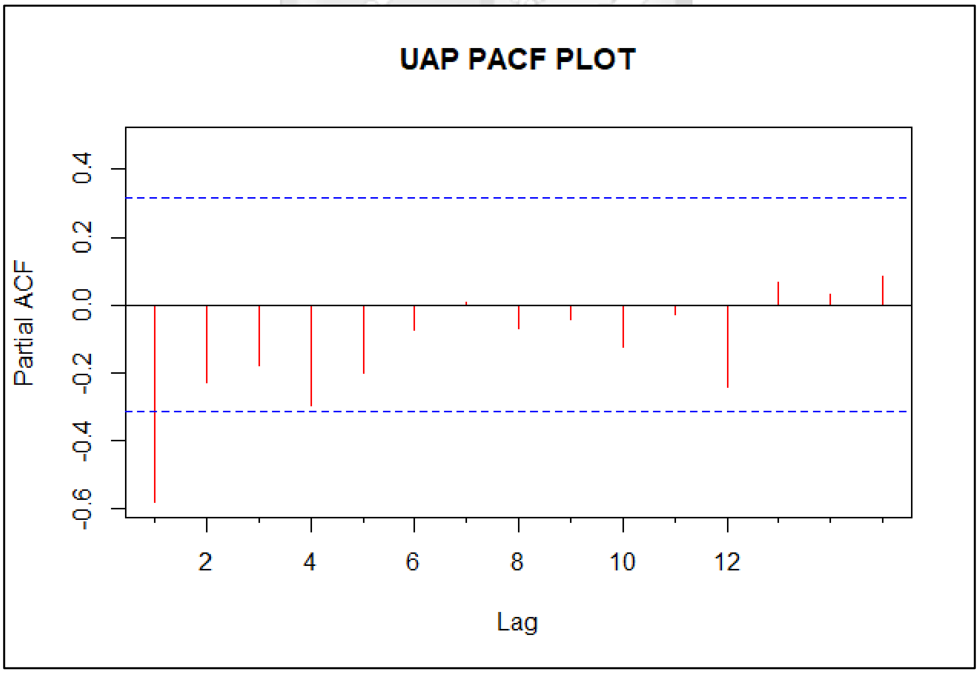
The Economic Times. (2024, May 15). *Insurance: What is General Insurance*. Retrieved from The Economic Times: https://economictimes.indiatimes.com/definition/general-insurance

Weindorfer, B. (2012). A practical guide to the use of the chain-ladder method for determining technical provisions for outstanding reported claims in non-life insurance. *Working Paper Series*, 1-17.

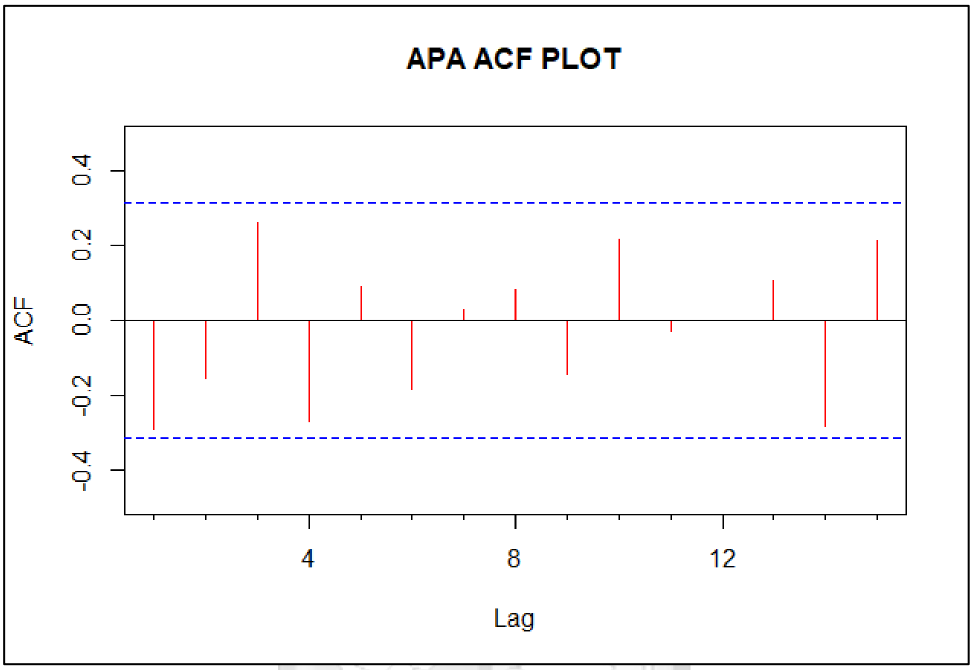
APPENDICES

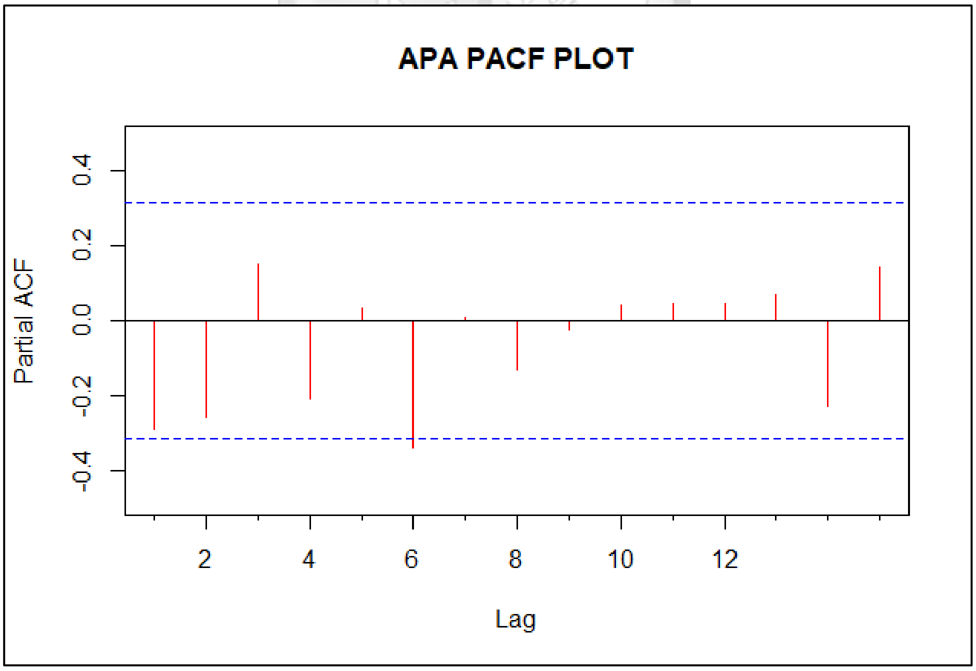
Appendix 1: The ACF and PACF plots for the differenced data for UAP.



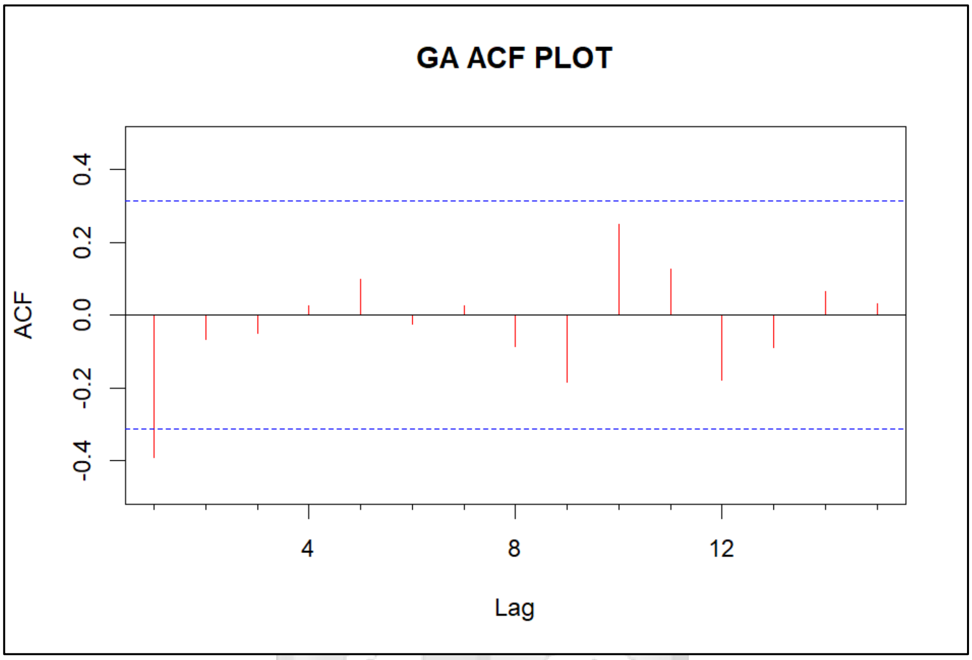


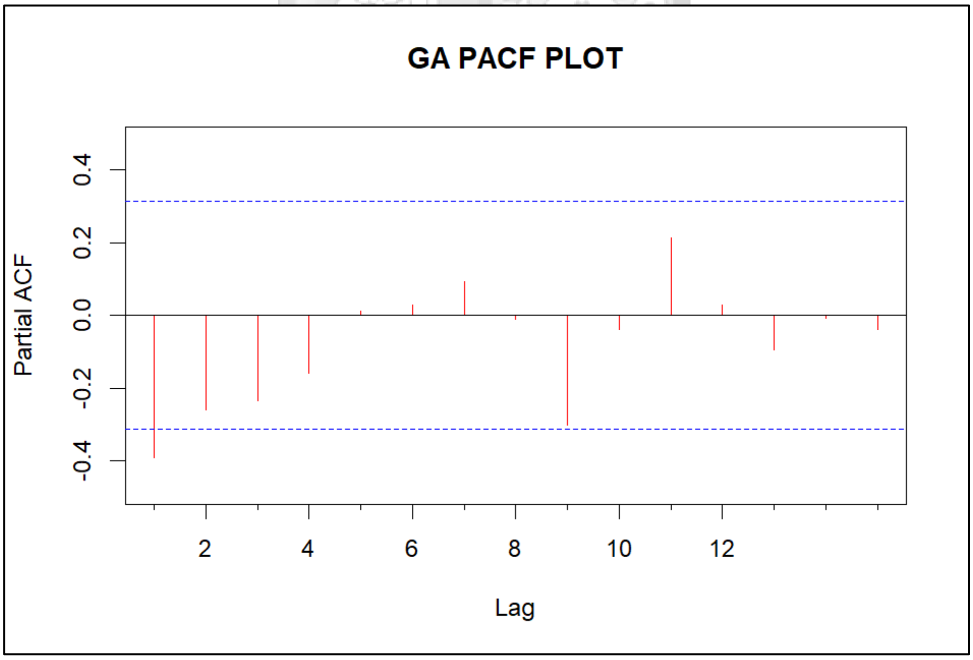
Appendix 2: The ACF and PACF plots for the differenced data for APA.



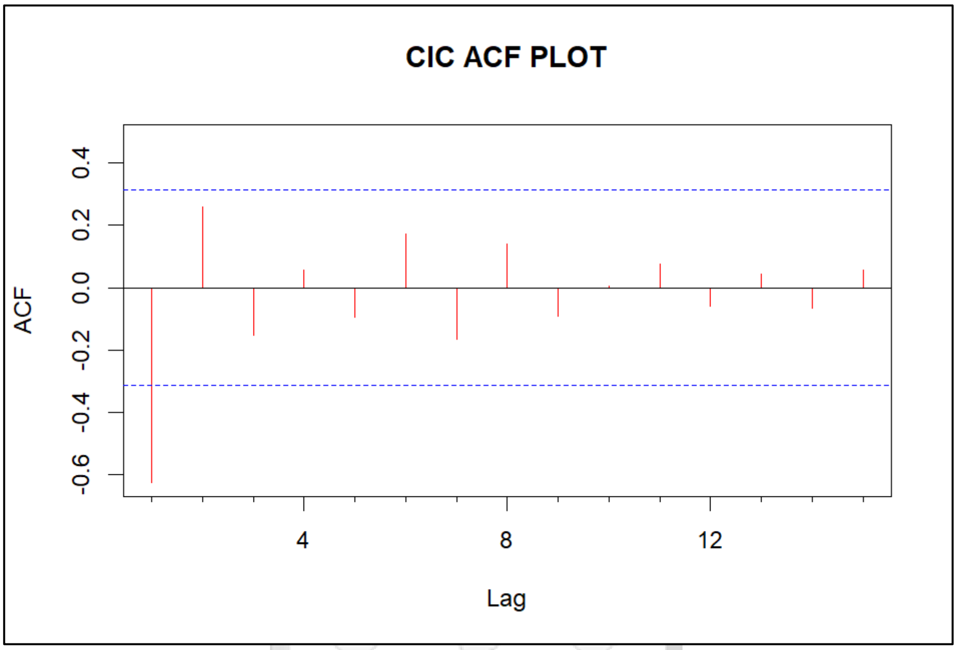


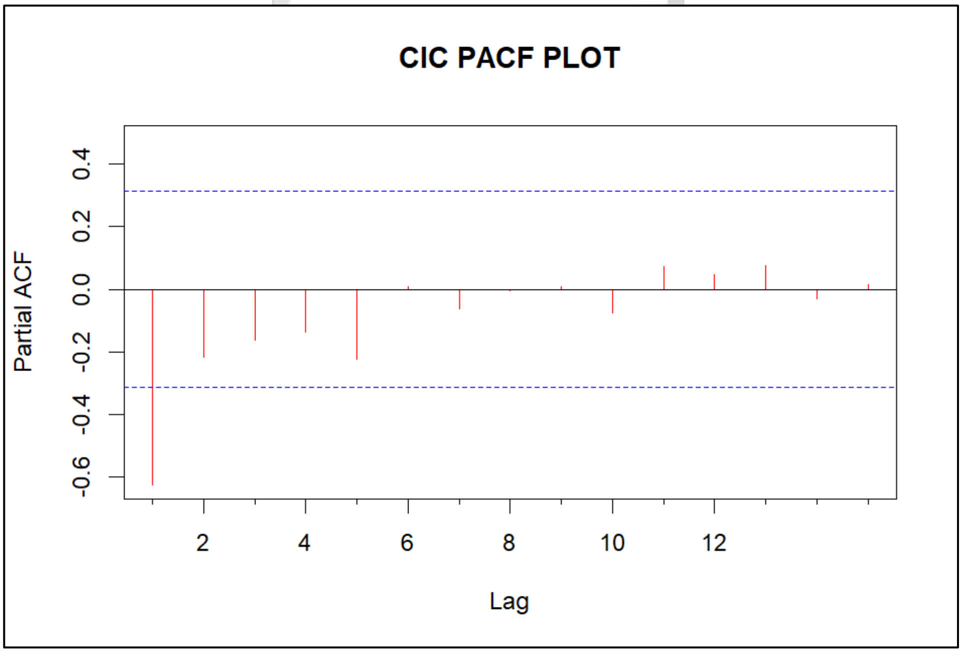
Appendix 3: The ACF and PACF plots for the differenced data for GA.





Appendix 4: The ACF and PACF plots for the differenced data for CIC.





Appendix 5: The ACF and PACF plots for the differenced data for Britam.