Prediction of Claim Fund Reserves in Insurance Companies Using the *ARIMA* Method

Goenawan Brotosaputro[1]*, Wiwin Windihastuty[2], Yohanes Setiawan Japriadi[3] , Rivai Ahsani[4]
Faculty of Information Technology [1], [3]
Atma Luhur Institute of Science and Business
Pangkalpinang, Indonesia

Faculty of Information Technology [2], [4]
Budi Luhur University
Jakarta, Indonesia
gbrotos@atmaluhur.ac.id[1], wiwin.windihastuty@budiluhur.ac.id[2], ysetiawanj@atmaluhur.ac.id[3],
2211600032@student.budiluhur.ac.id[4]

Abstract*— Insurance is a financial protection contract between a customer and an insurance company which is stated in the form of an insurance policy. Prediction of insurance claim reserve funds is necessary because the claim amount varies and the claim time can be the same. If at any time there is a claim that is so large that it exceeds the available claim reserve fund plus the claim occurs at the same time, it can cause the company to fail to pay the claim. This will certainly make the company's conduct decline, customer trust will be lost, and can cause the company to go bankrupt. The problem can be solved if the insurance company has sufficient claim fund reserves. Claim fund reserves are an important issue in insurance companies. This study aims to predict the claim fund reserves in insurance companies to anticipate varying claim amounts. Historical analysis of the value of claims with the ARIMA model approach is used to predict future claim values. We use claim value data that has been scaled in millions. 2020 to 2022 as training data and 2023 as test data. The Root Mean Square Error (RMSE) metric obtained is IDR 25,780.71; Mean Absolute Deviation (MAD) of IDR 14,421.89, and Mean Absolute Percentage Error (MAPE) of IDR 5,967.27; while the total actual claim value in 2023 is IDR 161,700.51 and the total predicted claim value is IDR 166,227.36; which means that an accuracy of 97% is obtained. The result of claim prediction value in one periodic year can give a favor to the management to make a decision, how much the claim funds should be prepared.

***Keywords*— prediction; classification; claim; ARIMA; insurance**

I. INTRODUCTION

In life, humans are not free from risks. Risk is any event that causes loss. Protection from various risks in everyday life has become a necessity. One institution that can provide this protection is an insurance company. Insurance is an agreement between the insurer and the insured to receive insurance premiums to replace or provide payment for losses experienced by the insured from an event.

Insurance promises protection to the insured against risks faced by individuals and organizations. Today, the insurance industry is experiencing very rapid growth. The insurance industry plays an increasingly important role in the Indonesian economy by providing benefits to both individuals and society. One of the characteristics of loss insurance is that the premium charged to policy holders depends on the number of claims and the size of the claims that have been submitted. The overall claim amount depends on the size of the claim and the frequency of claims in a particular period.

The number of claims submitted by policy holders represents the actual risk of the policy holder. Factors that cause the risk of damage or loss are disasters (perils) and dangers (hazards). This disaster factor can be influenced by the weather climate, which cannot be predicted or prevented. These factors will influence the number of claims.

Bankruptcy is the biggest risk that a company might experience. Technological developments, adaptability to the company's financial economic conditions are common factors causing bankruptcy. The company's financial condition has the largest percentage in influencing bankruptcy. To measure the level of risk and health of the company, an analysis of the company's financial statements needs to be carried out. From the analysis of the company's financial statements, it can be seen if the company is experiencing financial difficulties[1]. Bankruptcy is a significant event that can cause major losses to management, shareholders, employees, customers and the nation, so predicting company bankruptcy has become a hot topic for both industrial applications and academic research. In recent years, many studies have shown that machine learning techniques such as Artificial Neural Networks (ANN), Decision Tree (DT), Case Based Reasoning (CBR), Support Vector Machine (SVM) can be used as alternative methods for predicting company bankruptcy. Unlike statistical techniques, machine learning techniques do not assume a particular data distribution and automatically extract knowledge from training samples. Prediction performance depends on the details of the problem, the characteristics of the data structure used, the extent to which it is possible to separate classes using these characteristics, and the purpose of the classification[2]. Bankruptcy is a situation where a company experiences a shortage and insufficiency of funds to run or continue its

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2331, Copyright ©2025
Submitted : December 7, 2024, Revised : December 21,2024, Accepted : January 8, 2025, Published : January 31, 2025

business, a more serious consequence of bankruptcy is in the form of business closure or liquidation[3]. Bankruptcy prediction is very important for all organizations and agencies because it has a major impact on the economy and price increases will cause many social problems[4]. Several insurance companies in Indonesia have defaulted, namely the failure of the obligation to pay claims to the insured. Many factors cause insurance companies to default, one of which is due to poor calculation of claim fund reserves[5].

This research will be predict claim fund reserves using the *ARIMA* method so that insurance companies avoid bankruptcy due to inability to pay when claims occur that exceed the available claim funds. The Arima-based method often used in predicts the stock price in a company like said in [6] and have the trend predictions better than other machine learning methods such as linear regression, random forest, decision tree and gradient boosting machine [7]. Measurements should be carried out on research results to determine whether the implementation of a method is appropriate, such as in [8] which produces a smaller MAPE using ARIMA compared to using Single Exponential Smoothing in forecasting Indonesian cocoa export. Metrics other than MAPE are Mean Square Error (MSE) and Mean Absolute Error which are also used in [9] to predict dengue fever cases using ARIMA methods and have a good result.

The *Autoregressive Integrated Moving average (ARIMA)* method or commonly called the Box-Jenkins method is a method that was intensively developed by George EP Box and Gwilym Jenkins in the 1970s. The Box-Jenkins model group included in this method includes: *Autoregressive (AR), Moving average (MA), autoregressive-moving average* (*ARMA*), and *Autoregressive Integrated Moving average (ARIMA)*. The single value of the time series cannot be predicted, but if an analysis is carried out on the entire value of the time series, it has a certain pattern[10].

The data to be processed is data that has a record of the time the claim occurred and the value of the claim (numerical). The data obtained is analyzed to determine the pattern of past data that has been collected using the *ARIMA* model which can predict short-term forecasting for non-stationary data. Time series models such as *ARIMA* attempt to predict future conditions by using historical data and predicting the future. *ARIMA* prediction results in several studies are suitable for predicting variables that are very sensitive to short-term changes.

The *ARIMA* model uses dependent variables and ignores independent variables to produce accurate short-term forecasting. This research calculates predicted claim values based on previous claim values recorded in time series form. This research can provide input to insurance companies as a reference in predicting the provision of claim funds so that they do not cause problems in the future for their customers.

II. Methodology

The methodology in this research uses a quantitative approach whose aim is to calculate various things in an effort to explain what is observed and is useful for generalizations, predictions, and causal explanations. Another characteristic is that data is obtained in a structured manner in numerical form and in large quantities[11]. Previous research shows that ARIMA can be used to predict many things and get good results, so in this research the ARIMA method will also be used to predict the reserve funds of claims at insurance companies.

A. Research Flow

In predicting claims fund reserves using the *ARIMA* model, there are several stages that need to be carried out. This stage includes data collection, data preparation, data transformation, data sharing, *ARIMA* modeling, evaluation and measurement of prediction results which can be seen in Fig. 1.

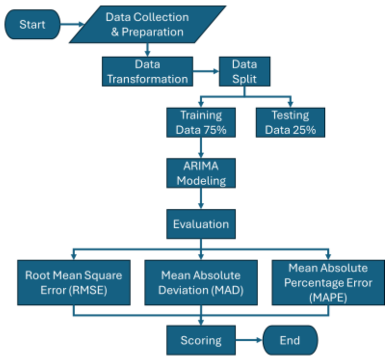


Fig. 1 Research Flow

B. Data Collection and Preparation

Claim value data for the analysis process was obtained from one of the leading insurance companies in Indonesia. The data obtained in Excel format is 8,301 rows with 25 columns which store claim data in a time series from 2019 to 2023. From the 25 existing fields, researchers only used the [Date of Loss, Currency and Gross Claim] fields with the aim of predicting claim value in the following times, so it is suspected that these 3 fields can be used to predict the value of the next claim which can ultimately predict the total claim fund reserves that must be provided by the insurance company. This raw data can be seen in Fig. 2. Data in 2019 that contained only 4 rows were eliminated because its not likely the others in 2020 until 2023.

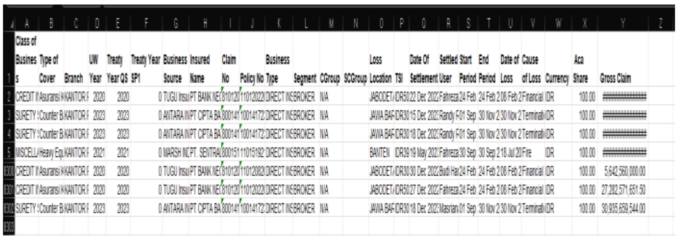


Fig. 2 Raw Data

C. Data Transformation

Based on observations and data analysis, we found that Currency consists of 5 different currencies. Standardization is needed for the differences in these currencies. Conversions were made from Australian Dollar, Euro, Singapore Dollar, and

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2331, Copyright ©2025
Submitted : December 7, 2024, Revised : December 21,2024, Accepted : January 8, 2025, Published : January 31, 2025

US Dollar to Rupiah. We used historical currency data by utilizing the site https://www.xe.com/ and to simplify it, the exchange rate was done by searching for the average value history from the 1st of each month to the end date of it. The value obtained was used as the Gross Claim exchange rate in the same month and year and then the claim value in foreign currency was converted to Indonesian Rupiah. After the conversion was completed, data aggregation was made by calculating the claim value per month per year which was represented in the occurrences column and the total claim per month as shown in TABLE 1.

Some of total claims per month have very large value, so to makes the good visibility of data, these value will be scaled in millions.

ARIMA has ability to make predictions with a single historical data, so in this research we only use the Total Claims per month and month as a time series. Time series in the form of dates require complete dates, so that for each month the 1st is given, the data are ready to be applied to the *ARIMA* model is presented in *TABLE* 2.

D. Data Split

To validate whether the model can be applied, the data needs to be divided into training data and test data. The training data using three years of data claim value per month from January to December 2020 until 2022 (75%) and test data using one year claim value per month from January to December 2023 (25%).

TABLE 1. Number of incidents and total claims per month

	2020		2021	
Month	**Occurences**	**Total Claim**	**Occurences**	**Total Claim**
Jan	10	433,189,538.48	197	6,955,649,415.77
Feb	1	17,181,818.00	247	100,992,041,856.33
Mar	3	23,986,616.35	156	5,786,845,736.36
Apr	11	3,184,744,171.47	255	11,591,699,592.74
May	15	177,464,989.69	136	3,508,105,175.11
Jun	20	2,737,806,501.20	128	2,218,888,646.71
Jul	57	6,971,070,261.27	132	1,164,053,047.63
Aug	27	38,516,815,629.18	142	4,736,044,421.67
Sep	99	3,331,185,777.17	155	5,545,840,503.78
Oct	78	3,461,902,091.75	254	11,126,383,570.43
Nov	54	1,361,843,072.62	204	9,597,759,035.99
Dec	77	2,237,629,991.97	226	10,087,082,882.02

	2022		2023	
Month	**Occurences**	**Total Claim**	**Occurences**	**Total Claim**
Jan	222	8,267,771,542.23	173	4,430,835,869.42
Feb	165	2,982,610,888.24	196	5,029,537,017.84
Mar	222	9,291,233,284.94	177	6,899,123,050.30
Apr	177	2,497,450,290.58	167	5,631,003,547.19
May	240	52,063,370,024.67	166	5,284,908,740.93
Jun	246	104,268,400,239.60	154	5,932,366,203.28
Jul	213	36,285,032,455.71	415	12,875,165,010.95
Aug	172	3,296,534,460.00	506	12,543,437,955.69
Sep	156	22,788,867,600.88	516	3,753,778,425.75
Oct	176	3,402,032,912.49	365	1,140,997,927.52
Nov	429	6,824,707,679.67	121	98,159,305,889.88
Dec	235	12,356,519,724.81	4	20,047,980.45

TABLE 2. Claim by Month and Year (in Million)

Month	2020	2021	2022	2023
Jan	433.1895385	6955.649416	8267.771542	4430.835869
Feb	17.181818	100992.0419	2982.610888	5029.537018
Mar	23.98661635	5786.845736	9291.233285	6899.12305
Apr	3184.744171	11591.69959	2497.450291	5631.003547
May	177.4649897	3508.105175	52063.37003	5284.908741
Jun	2737.806501	2218.888647	104268.4002	5932.366203
Jul	6971.070261	1164.053048	36285.03246	12875.16501
Aug	38516.81563	4736.044422	3296.53446	12543.43796
Sep	3331.185777	5545.840504	22788.8676	3753.778426
Oct	3461.902092	11126.38357	3402.032912	1140.997928
Nov	1361.843073	9597.759036	6824.70768	98159.30589
Dec	2237.629992	10087.08288	12356.51973	20.04798045

E. ARIMA Model

The components of *ARIMA*, as the name suggests, are (1) *Autoregressive*, (2) *Integrated*, and (3) *Moving average*. The stages for doing *ARIMA* are *Autoregressive (AR)*, *Moving average (MA)*, *Autoregressive Moving average (ARMA)*, and finally *Autoregressive Integrated Moving average (ARIMA)*.

Autoregressive (AR)

The *autoregressive* model is a stationary model of time series data where the observation value at time *t* is influenced by the previous observation value. This model uses the *AR* order (*p*) or *ARIMA* model *(p,0,0)* expressed in equation *(1)*.

$$Yt = \beta 1\ Yt\text{-}1 + \beta 2\ Y\ t\text{-}2 + \dots + \beta p\ Y\ t\text{-}p + et \qquad (1)$$

explanation:
Yt = The series values are stationary
βi = *i*-th *autoregressive* parameter
et = White Noise error value at time *t*

Independent variables are a series of values of similar variables in the last few *t* periods. While at is an error or residual unit that describes random disturbances that cannot be explained by the model. *Autoregressive* calculations can be done in the following process:

1) Determine the model that fits the time series.
2) Determine the value of the order p (determine the length of the equation formed)
3) Estimate the *autoregressive* coefficient values β1, β2, β3, βk

Moving average (MA)

The *Moving average (MA)* model shows observations at previously influenced times. The *moving average* is denoted in *MA (q)* or *ARIMA (0,0,q)* which is written in equation *(2).*

$$Yt = et - \beta 1\ et\text{-}1 - \beta 2\ et\text{-}2 - - \dots - \beta q\ et\text{-}q \qquad (2)$$

explanation:
Yt = Stationary series values
βi = *Moving average* Parameter

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2331, Copyright ©2025
Submitted : December 7, 2024, Revised : December 21,2024, Accepted : January 8, 2025, Published : January 31, 2025

et = White noise / error or residual unit

Autoregressive Moving average (ARMA)

The combination of the *Autoregressive (AR)* and *Moving average (MA)* models will form a new model, namely *ARMA (autoregressive moving average)* with the *ARMA* order *(p,q)*. The general form of the *ARMA* equation is a combination of the *AR* and *MA* equations which are denoted in equation *(3)*.

$$Y_t = \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \ldots + \beta_p Y_{t-p} + e_t - \beta_1 e_{t-1} - \beta_2 e_{t-2} - \ldots - \beta_q e_{t-q} \quad (3)$$

explanation:
Y_t = Stationary series values
β_i = *Moving average* Parameter
e_t = White noise / error or residual unit

ARMA modeling has a basic theory of correlation and stationarity. This means that *ARMA* can be used when the time series has formed a stationary graph, or does not form an upward or downward trend. However, if the time series data is not stationary, then a differentiation process is needed to change the data to become stationary before it can be processed through *ARMA*.

Autoregressive Integrated Moving average (ARIMA)

The *AR*, *MA*, and *ARMA* models use the assumption that the time series data produced is already stationary. In reality, time series data is more often non-stationary. If the data is not stationary, the method used to make the data stationary is differencing for non-stationary data in the mean and the transformation process for non-stationary data in the variance. The general form of the *ARIMA* model can be expressed in equation *(4)*.

$$Z_t = \emptyset_1 Z_{t-1} + \emptyset_2 Z_{t-2} + \cdots + \emptyset_p Z_{t-p} + a_t + \theta_1 a_{t-1} + \cdots + \theta_p a_{q-1} \quad (4)$$

explanation:
Z_t = data at time t, $t = 1,2,3, \ldots, n$
Z_{t-i} = data at time $t-i$, $i = 1,2,3, \ldots, p$
Z_{t-i} = error at period $t-i$, $i = 1,2,3, \ldots, q$
a_t = error at period t, $= 1,2,3, \ldots, n$
$\emptyset_0$ = *Autoregressive (AR)* model constants
$\emptyset_i$ = coefficient of Z_{t-i} in the *Autoregressive (AR)* model
θ_i = coefficient of a_{t-i} in the *Moving average (MA)* model

The general process for *ARIMA* modelling is: visualize the time series data, make the time series data stationary, plot the correlation and autocorrelation charts, construct the *ARIMA* model and use the model to make predictions. The establishment of the *ARIMA* model is that the time series are required to have stationarity. It can be judged intuitively by making the original sequence diagram, and the stationarity of the sequence can also be verified by the inspection method. If the sequence is not stable, one can use data transformation or difference to make the data stable [10].

1. Visualize the time series data

Time series data visualization is very important in *ARIMA* methods because it helps identify data patterns and trends, and ensures that the model built is accurate and relevant[12]. In addition, data visualization also helps identify the suitability of the *ARIMA* model to the existing data, thereby improving the quality of predictions.

From the historical data in *Fig. 3* it can be seen that in order to predict the claim value in the future, it is necessary to create a proper relationship between the claim value and time. The graphic image can be used to analyze data patterns and trends, whether stationary, non-stationary or seasonal and the hypothesis *H0* is taken: non-stationary data and *H1*: stationary data.

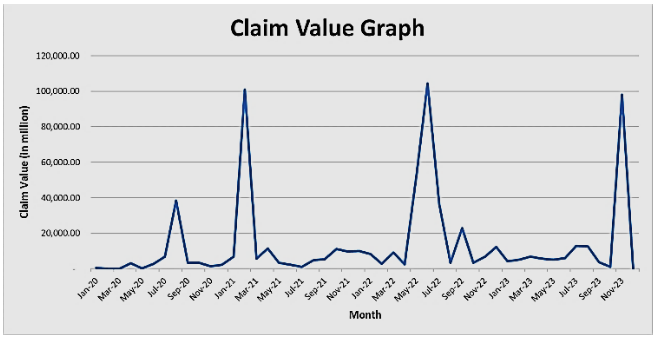
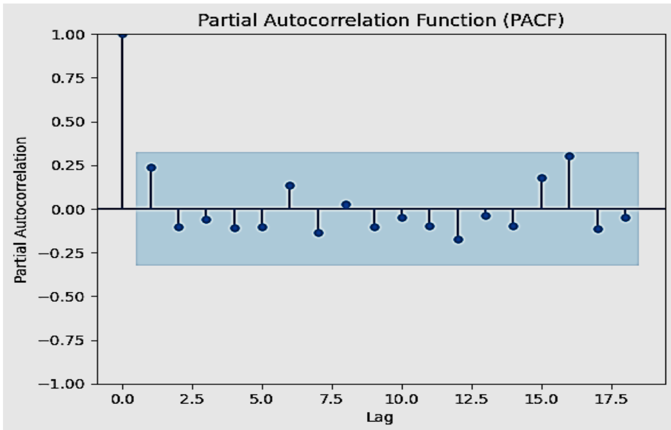


Fig. 3. Claim Value data visualization graph



2. Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) Plot Analysis

ACF and PACF are used to find the next parameters of *ARIMA*. Plot ACF (Autocorrelation Function) to determine the q value and PACF (Partial Autocorrelation Function) to determine the p value. In Fig. 5 it is known that the pattern decreases drastically after lag 1 and then the ACF value moves around zero. The appropriate model is MA(0) or the q value is 0 because it is not significant and there is no oscillation or cutting pattern at a certain lag. Because q is not significant, the value of q can also be 1. Meanwhile in Fig. 6 also shows a drastic decreasing pattern and then the PACF value moves around zero. The appropriate model is *AR*(1) or the p value is 1 which indicates a strong correlation between the current value and the previous value. The parameter p = 1, d = 0 and q

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2331, Copyright ©2025
Submitted : December 7, 2024, Revised : December 21,2024, Accepted : January 8, 2025, Published : January 31, 2025

= 0 will be used as the first parameter to model ARIMA(1,0,0) and the result will be evaluated to find the optimum parameters.

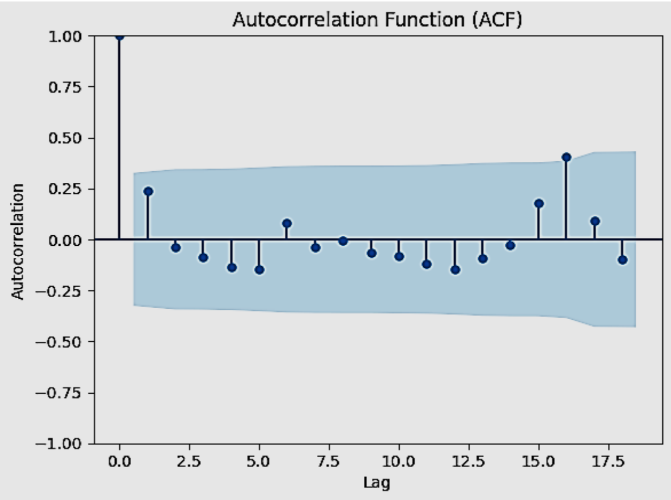


Fig. 5. Plot Autocorrelation (ACF)

F. Evaluation

After the values p, d, and q are obtained they will be used in the *ARIMA*(1,0,0) model. Based on training data and test data, training data from 2020 to 2023 was used and test data was data from 2023. The predicted claim value was obtained as in TABLE 3 and visualize in Fig. 4.

Based on the 2023 prediction results, several metrics were calculated, namely Root Mean Square Error (RMSE) of 25,780.71, Mean Absolute Deviation (MAD) of 14,421.89 and Mean Absolute Percentage Error (MAPE) of 5,967.27.

The ACF plot in *Fig. 5* shows that the residuals are randomly distributed, but there is an autocorrelation value that is outside the confidence interval limit (light blue area) at lag 16, which shows that the ARIMA(1,0,0) model is too simple, so it is necessary to modify the AR value or additional MA value to capture autocorrelation at lag 16.

Fig. 6. Plot Partial Autocorrelation (PACF)

TABLE 3. The 2023 Claim Fund Prediction with ARIMA(1,0,0)

Month	Claim	Forecast
Jan-23	4,430.84	13,532.02
Feb-23	5,029.54	13,807.24
Mar-23	6,899.12	13,871.68
Apr-23	5,631.00	13,886.77
May-23	5,284.91	13,890.30
Jun-23	5,932.37	13,891.13
Jul-23	12,875.17	13,891.32
Aug-23	12,543.44	13,891.37
Sep-23	3,753.78	13,891.38
Oct-23	1,141.00	13,891.38
Nov-23	98,159.31	13,891.38
Dec-23	20.05	13,891.38
Total	161,700.51	166,227.36

Experiments by changing the AR value or MA value several times obtained a fixed accuracy of 76% for predicting the value of claims accumulated over an annual period. However, if we look at the monthly period, at most 3 months the predicted value is close to the actual value as shown in *TABLE 4* and *Fig. 7* for January, February and August.

With the same metrics, the ARIMA(3,0,2) model obtains a Root Mean Square Error (RMSE) of 25,378.16, Mean Absolute Deviation (MAD) of 11,358.37 and Mean Absolute Percentage Error (MAPE) of 5,888.64.

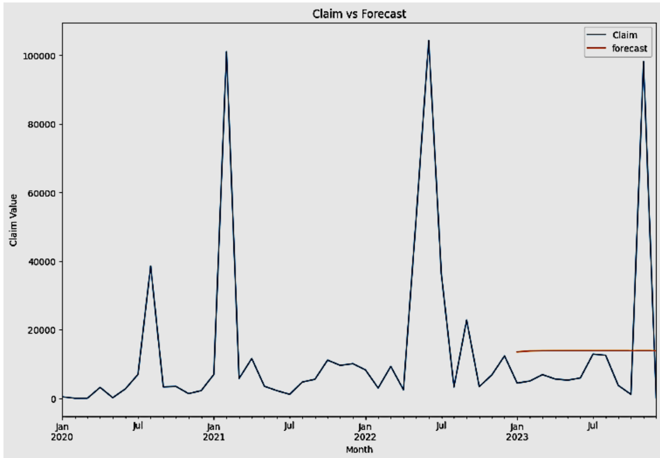


Fig. 4 Claim and Forecast comparison chart for 2023 with ARIMA(1,0,0)

TABLE 4. The 2023 Claim Fund Prediction with ARIMA(3,0,2)

Month	Claim	Forecast
Jan-23	4,430.84	5,275.16
Feb-23	5,029.54	5,796.02
Mar-23	6,899.12	5,898.88
Apr-23	5,631.00	8,541.00
May-23	5,284.91	9,022.58
Jun-23	5,932.37	11,161.42
Jul-23	12,875.17	11,169.29
Aug-23	12,543.44	12,694.75
Sep-23	3,753.78	12,358.17
Oct-23	1,141.00	13,477.99
Nov-23	98,159.31	12,976.87
Dec-23	20.05	13,851.67
Total	166,700.51	122,223.80

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2331, Copyright ©2025
Submitted : December 7, 2024, Revised : December 21,2024, Accepted : January 8, 2025, Published : January 31, 2025

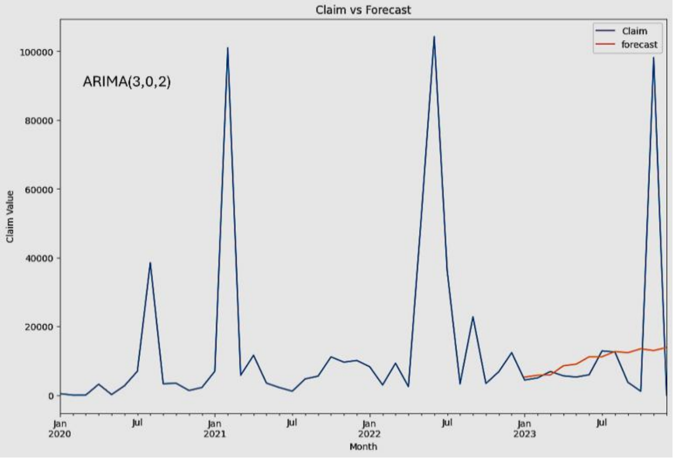


Fig. 7 Claim and Forecast comparison chart for 2023 with ARIMA(3,0,2)

Two experiments using ARIMA(1,0,0) and ARIMA(3,0,2) showed different prediction results for each month. The prediction value for each month in ARIMA(1,0,0) has almost the same value overall for 12 months, whereas in ARIMA(3,0,2) produces predictions with varying values in each month, especially in January, February and August whose value is close to the actual value.

III. Discussion And Results

Based on the Mean Absolute Percentage Error of ARIMA(1,0,0) model, a very large value is obtained, showing that the RMSE and MAPE values are greater than 50%, which means the model is inaccurate, this is due to the claim calculation which is calculated per month. However, if claims are calculated within 1 year, it can be calculated using TABLE 3 that the total actual claims in 2023 will be 161,700.51 and the total predicted claims in 2023 will be 166,227.36. This shows that if the insurance company will reserve funds for annual claims, this model can be used well, not for reserve monthly claim funds. The Mean Absolute Deviation (MAD) value of 14,421.89, compared with the actual average claim in 2023 which is 13,475.04 only has a difference of 946.85 or 7% also shows that this model suitable for predict the annual reserve claim funds.

Experiment by changing the parameters p, d and q becomes ARIMA(3,0,2) produce the metrics RMSE, MAD and MAPE becomes smaller than ARIMA(1,0,0), that are 25,378.16; 11,358.37 and 5,888.64 consecutively. This shows that the monthly prediction result better, but only for January, February and August. The total for one year shows the difference between actual and forecast claim is 39,476.71 or 24% deviated from the actual claim. Mean Absolute Deviation (MAD) has 11,358.37 which is deviated from the average actual claim 13,475.04 is 2,116.68 or 16%.

IV. Conclusion

Having sufficient claims reserves and predicting them accurately is an important step for insurance companies to manage their financial obligations, ensure business continuity, and fulfill obligations to policyholders. Without adequate claims reserves, insurance companies risk facing financial difficulties that can harm many parties.

The Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) metrics obtained from the *ARIMA*(1,0,0) model are greater than 50%, which means this model is not accurate, however, the Mean Absolute Deviation (MAD) metric The difference between the actual data was 7%, which means the model is very accurate. This contradiction is caused by the very fluctuating value of claims, especially in 2020 where the value was much smaller than the values in other years, this is thought to be due to the Covid-19 disaster at that time.

ARIMA(1,0,0) in the data studied shows better performance compared to ARIMA(3,0,2) which is shown by a difference of 3% in ARIMA(1,0,0) and 24% in ARIMA(3,2,0) for the total difference between predictions and actual values in one year, thus the parameters p = 1, d = 0 and q = 0 are suitable to be applied.

Comprehensively, based on the actual total claims in 2023 with the predicted total claims in 2023 using training data from 2020 to 2022, a difference of IDR 4,526.85 (in millions) is obtained or only 3% greater than the actual total claims or close to 97% accurate. This concludes that this model is suitable to be applied to predict the total value of claims annually, not monthly. The result of claim prediction value in one periodic year can give a favor to the management to make a decision, how much the claim funds should be prepared.

Acknowledgment

We express our deepest gratitude to the Institut Sains dan Bisnis Pangkalpinang and Universitas Budi Luhur Jakarta for providing the opportunity for their lecturers and students to implement *Tridarma Perguruan Tinggi* in the field of research. We also express special thanks to the Ministry of Education and Culture for the 2024 National Competitive Research Grant for Research on Affirmation Beginner Lecturers, who have funded this research until completion.

References

[1] A. W. Avrizal and E. Junianto, “Prediksi Kebangkrutan Perusahaan Menggunakan Algoritma Decision Tree Berbasis Teknik Resampling,” *eProsiding Tek. Inform.*, vol. 4, no. 1, pp. 101–109, 2023, [Online]. Available: https://eprosiding.ars.ac.id/index.php/pti/article/view/919/403.

[2] H. Saleh, “Prediksi Kebangkrutan Perusahaan Menggunakan Algoritma C4.5 Berbasis Forward Selection,” *Ilk. J. Ilm.*, vol. 9, no. 2, pp. 173–180, 2017, doi: 10.33096/ilkom.v9i2.97.173-180.

[3] S. Bahri and N. Widyawati, “Analisis Prediksi Kabangkrutan Pada Perusahaan Yang Di- Delisting Di BEI,” *J. Ilmu dan Ris. Manaj.*, vol. 4, no. 8, pp. 3–22, 2015, [Online]. Available: http://journal.stieputrabangsa.ac.id/index.php/jimmba/index.

[4] R. A. Nurdini, Y. Priyadi, and . N., “Analisis Prediksi Kebangkrutan Perusahaan Menggunakan Artificial Neural Network Pada Sektor Pertambangan Batubara,” *J. Sist. Inf. Bisnis*, vol. 8, no. 1, p. 107, 2018, doi: 10.21456/vol8iss1pp107-114.

[5] Y. Hikmah and I. R. Hikmah, “Perhitungan Cadangan Klaim Dengan Metode Chain Ladder Menggunakan Excel Dan Rstudio,” *MAp (Mathematics Appl. J.*, vol. 4, no. 2, pp. 122–131, 2022, doi: 10.15548/map.v4i2.4837.

[6] D. M. Putri and Aghsilni, “Estimasi Model Terbaik Untuk Peramalan Harga Saham PT. Polychem Indonesia Tbk Dengan Arima,” *MAp J.*, vol.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2331, Copyright ©2025
Submitted : December 7, 2024, Revised : December 21,2024, Accepted : January 8, 2025, Published : January 31, 2025

1, no. 1, pp. 1–12, 2019.

[7] M. Jia, H. Xu, and S. Zhang, *Comparison Analysis of ARIMA and Machine Learning Methods for Predicting Trend of US Semiconductor Stocks*, vol. 1. Atlantis Press International BV, 2022.

[8] M. S. Fahmuddin and S. M. Sri Mustika, "Perbandingan Metode ARIMA dan Single Exponential Smoothing dalam Peramalan Nilai Ekspor Kakao Indonesia," *VARIANSI J. Stat. Its Appl. Teach. Res.*, vol. 5, no. 3, pp. 163–176, 2023, doi: 10.35580/variansiunm193.

[9] E. Munarsih and I. Saluza, "Comparison of exponential smoothing method and autoregressive integrated moving average (ARIMA) method in predicting dengue fever cases in the city of Palembang," *J. Phys. Conf. Ser.*, vol. 1521, no. 3, 2020, doi: 10.1088/1742-6596/1521/3/032100.

[10] X. Zhang, Q. Zhou, S. Weng, and H. Zhang, "ARIMA Model-Based Fire Rescue Prediction," *Sci. Program.*, vol. 2021, 2021, doi: 10.1155/2021/3212138.

[11] S. Macdonald and N. Headlam, *Research methods handbook : introductory guide to research methods for social research*. 2008.

[12] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time series analysis: forecasting and control*. John Wiley and Sons Inc., 2015.

p-ISSN 2301-7988, e-ISSN 2581-0588
DOI : 10.32736/sisfokom.v14i1.2331, Copyright ©2025
Submitted : December 7, 2024, Revised : December 21,2024, Accepted : January 8, 2025, Published : January 31, 2025