Journal of
Actuarial Research
(JAR)

Forecasting Life Insurance Loss Reserves Using Time Series Analysis.

2016
actuarial
Science Club
2015

CARI
Journals

Journal of Actuarial Science
ISSN: 2958-0595 (Online)
Vol. 2, Issue No. 2, pp 21 - 31, 2024
www.carijournals.org

Forecasting Life Insurance Loss Reserves Using Time Series Analysis.

[1*]Angela Osei-Mainoo, [2]Daniel Logolam Atidebana, [3]Stephen Kwaku Nyarkoh

[1]Assistant Lecturer, Department of Statistics and Actuarial Science, C. K. Tedam University of Technology and Applied Sciences

[2]Department of Statistics and Actuarial Science, C. K. Tedam University of Technology and Applied Sciences.

[3]C. K. Tedam University of Technology and Applied Sciences,

http://orcid.org/0000-0001-8559-4888

Accepted: 13th Oct 2024, Received in Revised Form: 29th Oct, 2024, Published: 13h Nov, 2024

Abstract

Purpose: To determine an accurate predictive model that captures trends and seasonality in life insurance claims.

Methodology: Time Series Analysis is used to forecast Secondary data on life insurance claims. The researcher aims at using a non-actuarial and accurate predictive model that captures trends and seasonality in life insurance claims over time. The dataset, comprising monthly claim amounts from May 2018 to March 2024 obtained from an insurance company.

Findings: Life insurance claims increases with time hence the insurance companies are to set up an appropriate reserve to cater for future occurrence. It is realized that ARIMA (2,1,1) is appropriate for modelling the life insurance claim amounts with a least Log Likelihood value of -766.99, AIC value of 1521.97, AICc value of 1522.6, and BIC value of 1530.91. An ACF plot and Ljung Box test on the residuals shows the residuals are free from autocorrelation and free from heteroscedasticity respectively and hence the model is a white noise and adequate for further analysis. The result of the twelve months forecast indicates an increase in the life insurance claims.

Unique contribution to theory, practice and policy: Managers in the insurance companies should focus on risk management and reserve allocation of fund in order to meet short term and long term claims settlement.

Keywords: *Claim, Benefit, Insurance, Reserve, Loss*.

1. Introduction

Estimating life insurance loss reserves is essential for the long-term viability of insurance companies. Therefore, insurance providers are to maintain adequate reserves to meet future claims. In a middle income country like Ghana, where the insurance industry is growing appropriate actuarial methods should be used to ensure accurate loss reserve forecasting. This literature review explores the major techniques used for loss reserve forecasting, specifically, in the life insurance sector.

The National Insurance Commission (NIC), which supervises the insurance industry in Ghana, establishes the fundamental guidelines for determining loss reserves. The National Insurance Commission has a solvency requirement that every insurance company is supposed to meet before it is given the license to operate. (NIC, Ghana). The insurance sector in Ghana is confronted with highly volatile economic conditions, including exchange rate fluctuations, inflation, and currency devaluation. This makes forecasting with accuracy challenging (Sarpong, 2021).

These economic factors are compounded by limited access to high-quality data, making the application of advanced actuarial techniques challenging (Owusu-Ansah and Nyame, 2020).

As stated by Kuo et al. (2019), the insurance industry is gradually Utilizing cutting-edge technology and these tools enable a more precise and adaptable modeling by integrating extensive datasets with real-time information. In contrast, Ghana's life insurance sector still depends on traditional methods due to limitations in both technology access and data availability (Adu-Boahen, 2020). Nonetheless, there is a rising interest in implementing AI and machine learning solutions, especially in urban regions where improvements are being made to data infrastructure (Mantey and Owusu, 2021).

The ability to predict claim reserves is a problem faced by insurers and this has a great effect on the management and underwriting process. This becomes a serious challenge when the reserves are not appropriately estimated. In the event of a loss occurring within the coverage period, the insurance company compensates the policy holder (insured) by an amount which is referred to as the claim. However, the amount requested is the loss amount. Every business aims at optimizing resources hence there should be an adequate amount to cover future liabilities and this amount kept is called loss reserves. (Taha et al., 2021)

Loss reserves to the insurer is a loan owed to its clients. There are several techniques for modelling insurance loss reserves. Some of which are the classical actuarial chain ladder or the stochastic methods and machine learning based reserve prediction methods. (Avanzi et al.,2016).

According to Ibrahim et al., (2011), loss reserves enables the insurance companies to make strategic decisions such as underwriting decisions, investment decisions and corporate decisions.

Clark, 2023 explores the impact of emerging health trends on claims. Studies examine how advancements in medical treatment for chronic diseases (e.g., cancer) can lead to both increased life expectancy and potentially higher claim severity due to longer treatment periods. Additionally, the growing prevalence of mental health conditions is being investigated for its influence on claim rates, particularly for younger policyholders (Ettner, 2022).

Havnes and Zhang (2021) focused on the complex interplay between socioeconomic factors and life insurance claims. The Studies explores how income inequality, educational attainment, and access to healthcare can influence both the likelihood of having life insurance and the risk of filing a claim.

Wiatrowski et al., 2022 is of the view that lifestyle choices, such as substance abuse and risky leisure activities, has a great impact on risk assessment models.

Furthermore, Policy Lapse rate and Reinstatement can indirectly impact claim frequency. These behaviours are mainly due to economic downturns and changes in financial literacy (Borensztein and Ozdeser, 2019).

Several reasons account for the delay in claim settlements. Some of which are missing documentation, errors in policy information, or inconsistencies in beneficiary designations, concealment of material fact and so on. (Alay et al., 2018).

Also, Complex Claim Investigations, in cases of suspicious deaths, such as accidents or suicides within the exclusion period, insurers may conduct thorough investigations to rule out fraud. This can involve collaborating with law enforcement agencies and obtaining additional medical records, which can extend the settlement timeline (Kuo et al., 2019).

Wachter and Young, 2011 are of the view that Slow retrieval of Medical Records such as obtaining medical records from hospitals and healthcare providers can be time-consuming especially if the deceased received treatment from multiple facilities. Insurers rely on these records to assess the cause of death and determine if any policy exclusions apply.

Additionally, miscommunication between beneficiaries and insurers can lead to delays. This can occur if beneficiaries are unclear about the required documentation or if insurers fail to provide timely updates on the claim status. Internal Inefficiencies in the claim process procedures within insurance companies, such as manual paperwork handling inefficient information technology systems, can contribute to delays (Xu et al., 2021). Streamlining internal processes and implementing automation can significantly improve claim settlement times.

2. Method

ARIMA Model

Consider a process $\{y_t\}$ under ARIMA (p, d, q). If $(1-B)^d y_t$ is described by a stationary ARMA (p, q) model which is differenced using order d order to achieve stationarity. The order d is the order of Integration (I) of the series.

In general, the ARIMA model can be written as

$$\emptyset(\mathrm{B})(1-B)^d Y_t = \theta(\mathrm{L})\varepsilon_t \qquad (1)$$

$$\emptyset(B) = 1 - \emptyset_1 B - \emptyset_2 B^2 - \cdots - \emptyset_P B^P \tag{2}$$

$$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q \tag{3}$$

$\emptyset_t$ is the i^{th} autoregressive parameter, Θ_t is the i^{th} moving average parameter, p, q and d denote the autoregressive, moving average and differenced order parameter of the process respectively. Δ and B denote the difference and backward shift operators respectively.

Autoregressive process

A time series variable follows an Autoregressive (AR) process if the current value depends on its past values. That is the future of the series can be predicted using past values. For instance, statistically significant lag-1 autocorrelation indicates that the lagged value y_{t-1} might be useful in predicting y_t.

A simple model that makes use of such predictive power is

$$y_t = \emptyset_0 + \emptyset_1 y_{t-1} + \varepsilon_t \tag{4}$$

A general autoregressive process of p^{th} order has the form

$$y_t = \emptyset_0 + \emptyset_1 y_{t-1} + \emptyset_2 y_{t-2} + \cdots + \emptyset_p y_{t-p} + \varepsilon_t \tag{5}$$

The ε_t is assumed to be a white noise process .

Moving average process

Moving Average processes shows that the present value has something to do with the past residuals.

A general moving average process of q^{th} order has the form

$$y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} \tag{6}$$

With a constant mean μ, coefficients θ_1, and θ_2 and θ_q and an error term ϵ_t.

Autocorrelation function (ACF)

The correlation between y_t and its value at another time, say, $t + k$ is called the Autocorrelation at lag k . The population autocorrelation at lag k is defined as:

$$\rho_k = \frac{E[(y_t - \mu)(y_{t+k} - \mu)]}{\sqrt{E[(y_t - \mu)^2]E[(y_{t+k} - \mu)^2]}} = \frac{Cov\,(y_t, y_{t+k})}{Var((y_t)} = \frac{\gamma_k}{\gamma_0} \tag{7}$$

The collection of the values of $\rho\, k$ for k = 0,1,2, … is called the Autocorrelation Function (ACF) or Correlogram. The autocorrelation at lag zero is always 1, that is $\rho_0 = 1$.

Journal of Actuarial Science
ISSN: 2958-0595 (Online)
Vol. 2, Issue No. 2, pp 21 - 31, 2024
www.carijournals.org

Partial autocorrelation function (PACF)

According to the partial correlation coefficient in gross sectional regression, the partial autocorrelation coefficient π describes the supplementary information, which is provided by the additional lag. Thereby, the existing information derived from the previous lags, is taken into consideration.

However, if there exist more than one explanatory variable, as for example in the following model,

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \alpha_3 y_{t-3} + \mu_t \qquad (8)$$

Then one could ask, which additional influence the variable y_{t-2} or the variable y_{t-3} has, on the explanation of y_t . The partial autocorrelation of lag k is calculated by,

$$\pi_k = \frac{\rho s_k - \sum_{j=1}^{k-1} \pi_{k-1}, j . \rho_{k-j}}{1 - \sum_{j=1}^{k-1} \pi_{k-1}, \ j . \rho_{k-j}}, k > 1 \qquad (9)$$

3. Results and Discussion.

Table 1 Descriptive Statistics of Monthly Life Insurance Claims

Statistic	Mean	Median	Std dev	Max	Min	Skewness	Kurtosis
Values	352914	364236	168167	641553.26	58722.46	-0.08	-1.25

The descriptive statistics shows the minimum (58721.46), and maximum (641553.26) claim amounts experienced in May 2018 and February 2024 respectively. Also, the table shows 352914 as the mean and 364236 as the median, 168167 as the standard deviation. The Table also shows negative values of the skewness and kurtosis, 0.08 and 1.25 respectively, which indicates the data is tailed to the left and the distribution is platykurtic.

Stationarity of the data

H_0: The series is stationary

H_1: The series is non stationary

The Augmented Dickey fuller (ADF) test and the Kwiatkowski-Philips-Schmidt-Shin (KPSS) Test is used to test for the stationarity of the data. The data was differenced for the first time for stationarity and since p value 0.01 from the ADF test is less than 0.05, and the KPSS test value of 0.1 is greater than the p value 0.05 we have enough evidence against the null hypothesis. Therefore, we fail to reject the null hypothesis and conclude that the data is stationary.

Table 2 ADF and KPSS Test for the differenced data

ADF		KPSS	
Statistic	P-value	Statistic	P-value
-4.4359	0.01	0.035673	0.1

Journal of Actuarial Science
ISSN: 2958-0595 (Online)
Vol. 2, Issue No. 2, pp 21 - 31, 2024

The ACF plot of residuals and the p-values of the Ljung Box statistic were done to ensure the model was free from autocorrelation and conditional heteroscedasticity before the findings were obtained.

The Ljung Box test statistic is given by:

$$Q_m = n(n+2)\sum_{k=1}^{n}(n-k)^{-1} {r_k}^2 \approx {X_{m-r}}^2 \qquad (10)$$

Where ${r_k}^2$ = the reisduals autocorrelation at lag k

n= the number of residuals

m = the number of time lags included in the test

The model is considered adequate only for a large p value associated with Q**.**

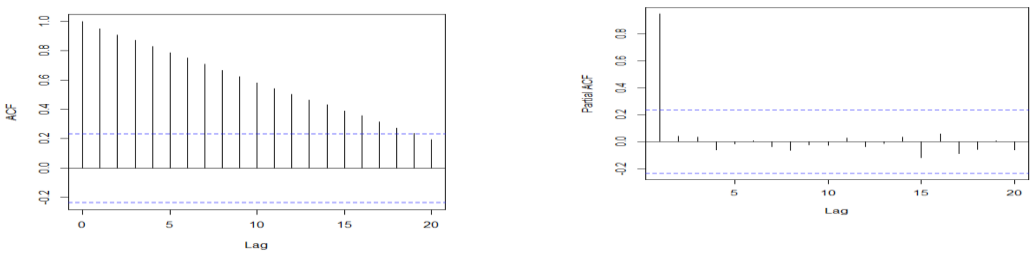

Figure 1 Plot of ACF and PACF for life insurance claim amounts

Figure 1 and figure 2 gives the ACF and PACF of the original life insurance claim amounts data. It can be observed that the autocorrelation functions show a significant slow decay nature in the spikes and the PACF have a very significant spike at lag 1. This is an indication of non-stationarity.

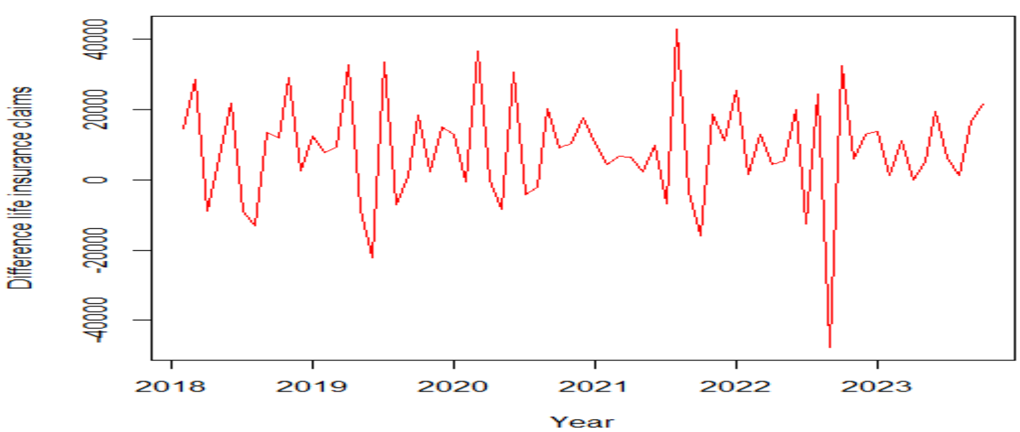

Figure 2 Difference life insurance claims

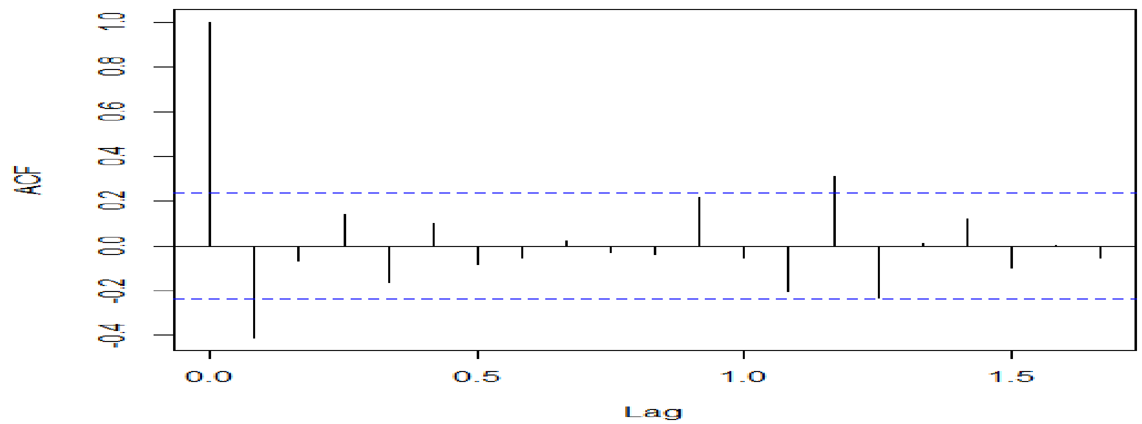

Figure 3 ACF plot of the Differenced data

Table 4 ARIMA model for Monthly Life Insurance Claims in Ghana Cedis.

ARIMA MODEL	AIC	AICc	BIC	Log Likelihood
ARIMA (1 1 1)	1531.83	1532.2	1538.54	-762.92
ARIMA (1 1 11)	1537.03	1543.65	1566.07	-755.52
ARIMA (1 1 14)	1525.35	1535.81	1561.1	-746.68
ARIMA (1 1 15)	1538.59	1550.59	1576.57	-752.29
ARIMA (2 1 1)	1521.97*	1522.6*	1530.91*	-766.99*
ARIMA (2 1 11)	1537.98	1545.76	1569.26	-754.99
ARIMA (2 1 14)	1521.21	1533.21	1559.19	-743.61
ARIMA (2 1 15)	1539.56	1553.24	1579.77	-751.78

***: means best model selected by the information criterion.**

The results in table 4 indicates that ARIMA (2 1 1) was considered the best model for the life insurance claim amounts from May 2018 to February 2023 because it has the least AIC value of 1521.97, AICc value of 1522.6, BIC value of 1530.91 and Log Likelihood value of -766.99.

Table 2 Parameter estimates of ARIMA (2, 1, 1) for monthly Life Insurance Claim in Ghana Cedis.

TYPE	Coefficient	SE	Z value	P-value
AR (1)	0.597109	0.110548	5.4013	6.615e-08*
AR (2)	0.402824	0.110553	3.6437	0.0002687*

Journal of Actuarial Science
ISSN: 2958-0595 (Online)
Vol. 2, Issue No. 2, pp 21 - 31, 2024
www.carijournals.org

MA (1)	-0.991827	0.016958	-58.4866	2.2e-16*

***: means significant at 5% significance level**

From the model it can be observed that the autoregressive components AR (1), AR (2) and the moving average component MA (1) are all significant predictors at 5% significant level.

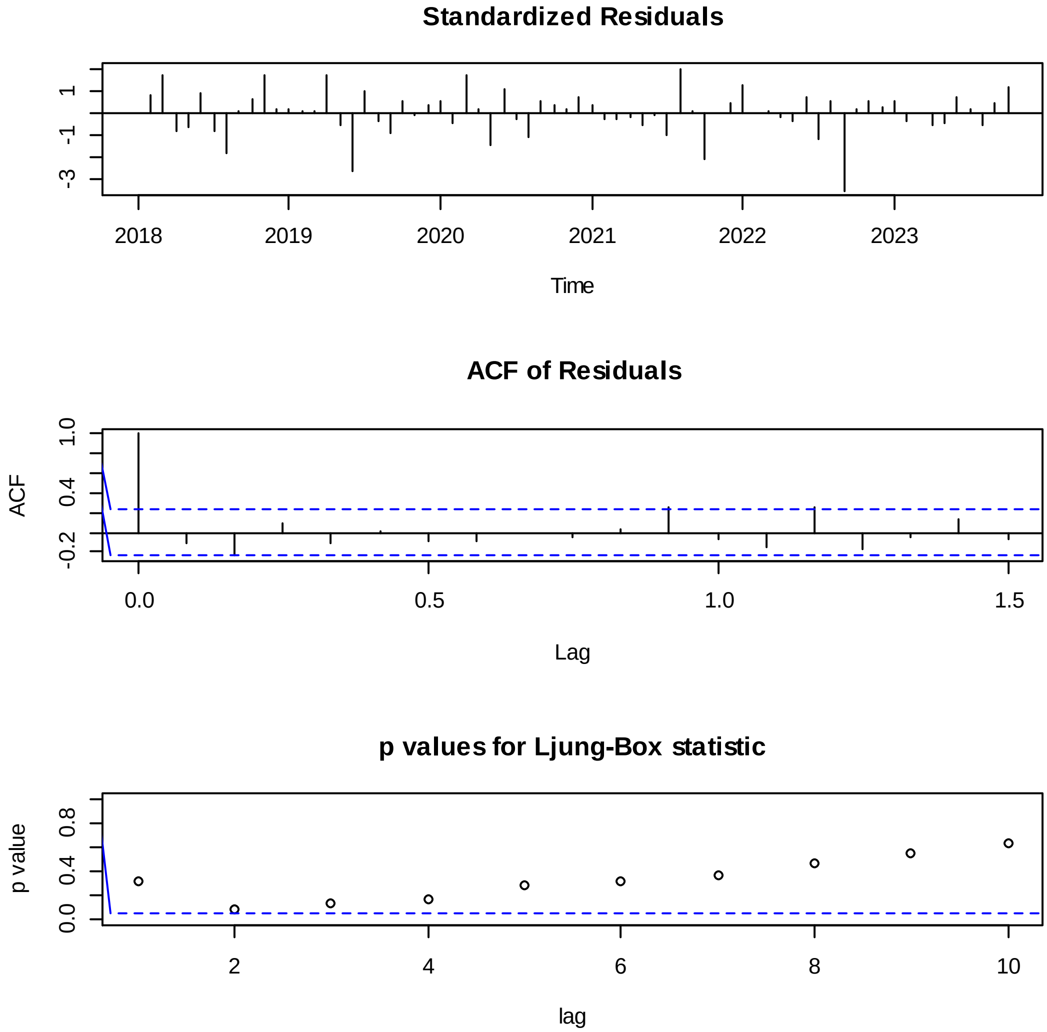

Figure 4 Diagnostic plot of Residuals of ARIMA (2,1,1)

To ensure that the fitted model, ARIMA (2,1,1) is adequate, model diagnostic tests were performed. The diagnostic test checks whether the residuals for the fitted model ARIMA (2,1,1) are white noise series.

From figure 7, the ACF plot shows that the residuals are free from autocorrelation even though there is a significant spike at lag 0. The points above 0.05 significant level in the p value for Ljung-

Journal of Actuarial Science
ISSN: 2958-0595 (Online)
Vol. 2, Issue No. 2, pp 21 - 31, 2024
www.carijournals.org

Box is an indication that the model is free from conditional heteroscedasticity. Hence, the model is white noise and adequate for further analysis.

Table 6 Forecast Life Insurance Claims

Month	Forecast	Lower Limit	Upper Limit
January	644378.6	616793.9	671963.2
February	654724.1	622375.0	687073.2
March	662039.6	623126.6	700952.5
April	670575.1	626759.1	714391.1
May	678618.6	629963.4	727273.8
June	686859.8	633804.4	739915.2
July	695020.8	637751.8	752289.8
August	703213.6	641935.9	764491.3
September	711393.0	646251.2	776534.7
October	719577.2	650699.9	788454.5
November	727758.9	655251.0	800266.9
December	735941.1	659893.4	811988.9

Table 6 shows the twelve (12) month's forecast of the life insurance claim amount using the ARIMA (2, 1, 1) at 95% confidence interval.

4. Conclusion and Recommendation

This study was carried out to forecast the trend of life insurance claims and to develop ARIMA model for the life insurance claim amounts. From the descriptive statistics, the skewness value of -0.08 shows the data is skewed to the left of the normal distribution (left tail) and the distribution is platykurtic based on the kurtosis value of -1.25.

Stationarity test was done on the data using the Augmented Dicky Fuller (ADF) and KPSS test, the analysis showed that the original data was non-stationary hence the data was differenced once for stationarity.

ARIMA (2,1,1) is considered the best model for the life insurance claim amount from May 2018 to February 2023 because it has the least AIC, AICc, BIC and Log Likelihood values as shown in table 4.

From figure 7, the ACF plot shows that the residuals are free from autocorrelation even though there is a significant spike at lag 0. The points above 0.05 significant level in the p-values for Ljung

Box Statistic is an indication that the model is free from conditional heteroscedasticity. Hence the model is white noise and adequate for further analysis.

Therefore, ARIMA (2,1,1) is used to forecast the life insurance claim amounts for the next twelve (12) months as shown in table 4.5. The forecast depicts an increasing trend in the claim amount.

To ensure liquidity and financial stability of the company, there should be emphasis on risk management and reserve allocation, the model predicts a rise in life insurance claims, therefore the insurance company should proactively increase their reserves to cover these anticipated payouts.

Furthermore, management is supposed to ensure pricing adjustments, given an upward trend in claims for the next twelve months, it may be necessary to review and adjust the premium pricing strategies. This will help maintain profitability in the face of rising payouts.

Finally, managers should enforce customer trust by informing policyholders on potential premium adjustments and changes in policies.

References

Adu-Boahen, J. (2020). The future of actuarial science in Ghana: A review of technological advancements. Ghanaian Journal of Insurance Studies, 14(2), 33-45.

Avanzi B.,Taylor G.,Vu P.A. and Wang B.(2016). Stochastic Loss Reserving with Dependence: A flexible multivariate tweedie approach "Insurance, Mathematics and Economics.,71,63-78.

Clarke S.R(2023). A treatise on the of Insurance. Books on Demand.

Frees E. (2018) Loss data analytics arXiv,1-319.

Islam,M.R.,Liu,S.,Biddle ,R.,Razzak,I.Wang ,X.,Tilocca,P.,and Xu,G.,(2021) Discovering dynamic adverse behavior of policy holders in the life Insurance Industry.Technological Forecasting and Social Change,163,120486.

Kuo, K., (2019). Generative synthesis of insurance datasets. ArXiv preprint arXiv:1912.02423.

Kuo, W., Huang, Y., & Tsai, M. (2019). Machine learning applications in actuarial science: Trends and future outlooks. Journal of Insurance Technology, 29(1), 9-27.

Mantey, K., & Owusu, G. (2021). AI and machine learning in the insurance industry: Prospects for Ghana. Ghana Technology Review, 17(2), 55-71.

National Insurance Commission (NIC). (2022). Insurance Industry Report. Accra: NIC Publications.

Owusu-Ansah, K., & Nyame, J. (2020). The role of economic factors in life insurance reserving: Evidence from Ghana. West African Journal of Insurance Economics, 8(2), 47-62.

Sarpong, J. (2021). Regulatory challenges in Ghana's insurance sector: The way forward. African Journal of Insurance Policy, 13(1), 61-77

Taha A., Cosgrave B., Rashwan W., and Mckeever S. (2021). Insurance Reserve Prediction Opportunities and Challenges, "the proceeding of international conference on Computational Science and Computational Intelligence (CSCI),2021, Las Vegas, USA.DOI:10.214/trtz8.t14

Taha T.A.E.A., Ibrahim Y .and Sobri M.M. (2011) Forecasting General Insurance Loss Reserves in Egypt. African Journal of Business Management 5(22), 8961-8970.

Wachter, T.V., Song, J.and Manchester. (2011). Trends in employment and earnings of allowed and rejected applicants to social security disability insurance programme. American Economic Review,101(7),3308-3329.

Yousaf I. Jareno F. and Martinez-Serna, M.I. (2023). Extreme spillovers between Insurance tokens and Insurance Stocks:Evidence from the quantile Connectedness approach. Journal of Behavioral and Experimental Finance,39,100823.