

Modeling and Predicting of Motor Insurance Claim Amount using Artificial Neural Network

V SelvaKumar , Dipak Kumar Satpathi, P. T. V. Praveen Kumar, V.V. Haragopal



Abstract: In India the insurance industry is in its growth stage. It consists of 58 insurance companies of which 24 in life and 34 are non-life insurance. The Non-life Insurance companies which cater to motor insurance business presently utilize different trend models to forecast paid claim amount. Motor Insurance Claim amount prediction is one of the most difficult tasks to accomplish in financial forecasting due to the complex nature of data points. The main objective of this study is to determine a reliable time series forecasting model to predict own damage (OD) claim amount of motor insurance data in India from 1981 to 2016. In this context, the annual time series claim data was collected and modeled by using the Generalized linear model (GLM), Autoregressive Integrated Moving Average (ARIMA) and Artificial Neural Network (ANN) method. The validation of the model has been done by comparison of predicted and actual values for the period of 36 years. Also, different types of possible models were evaluated using Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC), Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE) for accuracy. The results showed that ANN outperformed other traditional time series models (GLM & ARIMA) for predicting the future own damage claim amount with a lesser residual error. Further, the outcome of these data analytics studies would help Insurance companies to have an idea about the expected future claim amounts with more accuracy. Thus, predicting the Motor insurance's own damage claim will help insurance companies to budget their future revenue.

Keywords: Stationary Process, GLM, ARIMA, Neural Network, TRAINLM, RMSE.

Manuscript received on March 15, 2020.

Revised Manuscript received on March 24, 2020.

Manuscript published on March 30, 2020.

* Correspondence Author

V SelvaKumar*, Research Scholar, Department of Mathematics, Birla Institute of Technology and Science-Pilani, Hyderabad, Telangana, India.

Assistant Professor, Bhavan's Vivekananda College of Science, Humanities and Commerce, Hyderabad. Email: 20170106@hyderabad.bits-pilani.ac.in.

Dipak Kumar Satpathi, Associate Professor, Department of Mathematics, Birla Institute of Technology and Science-Pilani, Hyderabad, Telangana, India. Email: dipak@hyderabad.bits-pilani.ac.in.

P. T. V. Praveen Kumar, Assistant Professor, Department of Mathematics, Birla Institute of Technology and Science-Pilani, Hyderabad, Telangana, India. Email: praveen@hyderabad.bits-pilani.ac.in.

V.V. Haragopal, Visiting Professor, Department of Mathematics, Birla Institute of Technology and Science-Pilani, Hyderabad, Telangana, India. Email: haragopal_vajjha@hyderabad.bits-pilani.ac.in

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

I. INTRODUCTION

In the past few years, the motor insurance sector in India has heavily suffered based on motor own damage claim from a commercial motor vehicle concerning the viability of the insurance companies and also due to promoters continuously infusing additional capital to maintain solvency. Also, the long-term claims intimation, the trend of courts in awarding more desirable reimbursement, non-cooperation of insured in declaring claim settlement and non-flexibility of courts in claim dealing with, are some of the issues faced by the industry. Model evaluation and Forecasting the time series statistics is essential in fields like Insurance, Finance, and Agriculture, etc.

The previous research on modeling of Insurance claim is very scanty, and few authors have considered the ARIMA model for prediction with respect to the fire insurance loss, property damage claim and future health care insurance cost of inpatient diagnosis. In the literature ARIMA modeling for different general insurance areas like fire insurance, property damage insurance, and health care insurance with different parameters with overfitting was performed. However, forecasting the nature of the exact time-series claim amount of data is difficult. Though, there are various forecasting techniques applied in different sectors, only the ARIMA method can be employed for forecasting claim size. Also, there has been no substantial amount of research work done concerning the Motor Insurance Claim amount in India. This remains a motivating factor for us to evolve and develop predictive models as compared with other studies. Therefore, the present study focused on evaluating different forecasting models viz., GLM, ARIMA, and ANN, which represents the linear and Artificial intelligence was explored for own damage claim amount. The objective is to develop an appropriate mathematical model that helps in forecasting future predictions based on the chosen characteristic of the data [1].

II. REVIEW CRITERIA

In this section, we summarize various works listed in the literature. David Cummins et al. [2] studies indicate that the econometric model produces better forecasts as compared to the ARIMA Model concerning Property damage Claim of the automobile insurance companies in the United States. However, the study based on forecasting future health care insurance cost of in-patient diagnoses, by applying various forecasting techniques, the ARIMA model provide a good model fit with a high degree of accuracy in predicting future costs [3].

According to another studies on Republic of Macedonia life insurance data, by applying two competent models i.e. Combined ARIMA and Moving Average Model, all the results showed that the combined ARIMA model gives the best fit to the original series [4]. Another study of forecasting Egyptian fire insurance loss ratio in Protection and

Indemnity Insurance Companies for the period 1972 to 2001 found that this method is not reliable for forecasting marine cargo [5]. Based on the study of Misr Insurance company in Egypt, the fire insurance segment revealed that the ARMA (1,1) model is a good prediction for loss ratio [6].

Another research article published in this regard discusses that the forecasted ARIMA (6, 1, 6) with overfitting for the Inflation in Ghana over 12 horizons [7]. But another study revealed that ARIMA (0, 1, 1) model is most suitable for predicting the inflation in Burundi [8]. Artificial Neural Network (ANN) and ARIMA models are most popular for forecasting the linear and non-linear time-series data. But, according to a study [9] proposed a hybrid method which gives the best forecasting results for the time series data compared with ANN, Zhang's hybrid and ARIMA models. According to Olayiwola Olaniyi Matthew et al. [10], the overfitted ARIMA (8, 1, 2) model forecasted stock price ranges from 138.66 to 141.49 for Nigerian Breweries Plc. A recent study concluded that people's reading patterns of the newspaper changed to digital media due to some negative impact on the printed newspaper demand and inaccuracy of supply. ARIMA (1, 1, 0) model predicted printed newspaper demand accurately and restrained the oversupply [11]. Based on the literature survey on various prediction models proposed, modeled, and suggested by various authors earlier, the ARIMA model is widely used for forecasting the time series data with accuracy and prediction [12].

Another research article [13], in the food manufacturing industry, showed that the ARIMA (1, 0, 1) model could be reliable to predict future demand. Madhavi Latha et al. [14] concluded that more uncertainty had been found when the forecast period is long term period, while less certainty exists in the case of short-term period of Bombay stock exchange return. According to another paper [15], ARIMA (5, 0, 1) (2, 0, 0) gives the best fit as per their claim to predict the average monthly rainfall in Rayalaseema district. Mohamed Reda Abonazel et al. [16] analyzed Egypt's annual GDP time series data for predicting the GDP for the years 1965 to 2016. ARIMA (1, 2, 1) model was used to predict the GDP for ten years. Another Researcher concluded that the seasonal ARIMA (1, 1, 0) model was used to forecast the International tourist visit to Bhutan with 91% forecasting accuracy [17]. In a recent study about forecasting solar radiation, ARIMA (1, 1, 2) model predicted the daily radiation, while the monthly radiation was predicted by ARIMA (4, 1, 1). The results revealed that the expected average monthly solar radiation ranges from 176 to 377 Wh / m² [18].

From the above discussions, it is observed that many authors have successfully fitted the ARIMA model for their studies. Some researchers showed overfitting for which the predictions may not be accurate as compared with the correct fitting of the model. However, the ARIMA model is applied to linear data function only. But it is uncertain to predict the exact nature of the data series, which has a non-linear

structure of the data points. We apply the artificial neural networks (ANN) to the non-linear data. This model is self-adaptive, data-driven, and non-linear, which has few assumptions. Several works have been done in time series analysis using traditional time series methods and ANN's from different fields. Raid et al. [19] study indicates that the multilayer perceptron neural network method is more suitable to predict river runoff as compared to the traditional regression model. Suppachai Pathunakul et al. [20] have applied the ANN model to predict the production rate and dust level enhances the mill's capability. Another study has estimated the market share of Thai rice in the global market by using ANN [21].

According to Zhu Ying et al. [22] have applied ANN model to predict market demand. Karin [23] compared three methods: ARIMA, Support Vector Machines (SVM), and ANN, to forecast the demand for consumer products in Thailand. Another study related to forecasting the electricity demand; concluded that the ANN model is superior to other traditional models [24]. Slimani et al. [25] also applied the ANN method to forecast the traffic flow by using Multilayer perceptron architecture. Q Yang et al. [26] have applied ANN modeling for achieving high academic performance within academic institutions. Thus, keeping in this view we have explored the data by considering the GLM, ARIMA model, and ANN model for predicting the own damage claims with accuracy and a better model fit.

III. RESEARCH DESIGN & METHODOLOGY

A. Objective of the Study

This research study aims to develop a time series model to forecast their own damage (OD) claim amount based on the historical sample data. To arrive at this, we have used consistent historical claim data from 1981 to 2016. The major contributions of the study are as follows:

1. To forecast the Own Damage Claim amount of insurance companies using GLM, ARIMA, and ANN.
2. To identify the most suitable forecasting technique among the three above mentioned model by considering the adequacy, suitability, and accuracy of the data modeling.

B. Data used for Research

Data used for the study is collected from different Insurance companies of India for distinctive 36 years from 1981 to 2016. The original data is of 108 column variables distributed over 50,000 rows (i.e., approximately fifty-four lakh data points). Out of this data, the variable analyzed is the OD claim variable, which is considered for fitting Generalized Linear Model, ARIMA, and Neural Network Model for predicting the future claim values by using MATLAB R2019b version.

C. Methodology:

Time-series forecasting models

To build a time series model based on a claim paid data, we evaluated the traditional statistical models, including GLM and ARIMA.

These models forecast the future values of a time-series data based on the linear function of past values and white noise term.

Generalized Linear Models (GLM):

Generalized Linear Model is the forecasting technique generated by a simple linear regression model. GLM is modeled as a linear relationship with an error term:

$$Y_i = \beta_0 + \beta_1 X_i + e_i \tag{1}$$

Where X_i is a time variable in a yearly unit; Y_i is an Own damage claims amount, e_i is a residual error term

Autoregressive Integrated Moving Average ARIMA (p, d, q):

In order to fit a model to the data efficiently, the model needs to be a stationary series. So, the first task is to ensure that the claim data be stationary series or not. Initially, the variable under consideration is not in the stationary form; hence, the trends and cyclic behavior have to be studied. To bring the data into stationary processes, the Autocorrelation function (ACF) and Partial autocorrelation function (PACF) are evaluated. These suggest the most applicable model for the data. Hence, the data is evaluated by the ARMA (p, q) model, which is the combination of the Autoregressive and Moving Average model. After studying the stationarity, we have considered the ARIMA.

Box – Jenkins model comprises a process approach to follow and repeat the process until attaining a high degree of contentment about the model and having reduced errors. Consequently, without any hindrance, we can forecast the variable under consideration.

The mathematical equation of Box and Jenkins [27], ARIMA (p, d, q) process can be expressed as AR (p) I (d) MA (q) in terms of the backward shift operator [28].

$$(1 - \alpha_1 B_1 - \alpha_p B_p)(X_t - \mu) = (1 + \beta_1 B_1 + \dots + \beta_q B_q) e_t \tag{2}$$

To identify the best ARIMA (p, d, q) model among several experiments performed, the following primary steps evaluated for the OD Claim variable:

- * Identification of a model (Plotting the ACF and PACF)
- * Adjusted R squared (larger the value, the better the fitting)
- * Relatively AIC, BIC, and RMSE are small
- * Estimation of parameters within the recognized model
- * Diagnostic Checks and Forecasting Also, by comparing more than one model, the least AIC and BIC are generally considered to be best for the fitted model.

Also, by comparing more than one ARIMA Model, the least AIC, BIC and RMSE is generally considered to be best for the fitted model. The other methodology implemented is ANN.

Artificial Neural Network (ANN):

Artificial Neural Network (ANN) is an information processing paradigm that is composed of a large number of highly interconnected processing elements (neurons) working together to solve specific problems. The signals are transmitted through a large number of nodes. The nodes

possess an associated weight, which is multiplied along with the incoming signal for any neural net. By applying the activation function to the net input, the output signals are obtained. The basic building characteristics of the ANN are Network structure, Training (setting the weights) and Activation function. The multilayer perceptron (MLP) neural network model is mostly used for forecasting the time series own damage claim amount. The MLP consists of one input layer, one or more hidden layers, and an output layer.

The input variable X_i , $i=1$ to n to the neuron are multiplied by weights W_{ki} . The output of a neuron is the input to the activation function. The mathematical equation of the multilayer perceptron feed-forward neural network model can be written as

$$y_t = w_0 + \sum_{i=1}^m w_i f(\sum_{j=1}^n w_{ij} y_{t-1} + w_{0j}) + e_i \tag{3}$$

Here ‘n’ refers to the number of Input nodes, ‘m’ is the number of hidden nodes; ‘f’ is the sigmoidal activation functions; ‘ α_i ’ is a weight from the hidden to output nodes, and. ‘ w_{ij} ’ are weights from the input to hidden nodes. The widely used activation function is logistic sigmoid and tangent hyperbolic.

The sigmoidal function can be expressed mathematically as follows:

$$f(x) = \frac{1}{1 + e^{-ax}} \tag{4}$$

In recent studies in deep learning, the suggested alternative sigmoid functions are Gaussian, Rectified linear, SoftMax and hyperbolic tangent function. The sigmoid hyperbolic tangent function can be defined as

$$\tanh(x) = \frac{1 - e^{-x}}{1 + e^{-x}} \tag{5}$$

The computation of Root Mean Square Error (RMSE) and Mean Absolute Percentage (MAPE) are shown in equations (3) and (4), are also computed.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \tag{6}$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| 100 \frac{y_i - \hat{y}_i}{y_i} \right| \tag{7}$$

IV. RESULT AND DISCUSSION

In this section, the discussion of our results based on the performance measures of various forecasting models GLM, ARIMA, and ANN are developed for the OD Claim amount variable.

A. The results of the ARIMA Model

Firstly, the data which is in non-stationarity is brought down to stationarity by differencing the process. Thus, the original process can then be reconstructed by integrating the differenced series, and the original series will be modeled by an Autoregressive Integrated Moving Average (p, d, q) [29].



The Augmented Dickey-Fuller (ADF) unit root test is significant ($p = 0.00001 < 0.01$), and we conclude that the data evaluated is stationary. Now, we computed the optimal lag length for the autoregressive parameter (p) and moving average parameter (q) based on autocorrelation function (ACF) and partial autocorrelation function (PACF) [30].

From TABLE I below, all best possible feasible ARIMA models are estimated, and numerous coefficients are analyzed.

The best estimated probable six models are compared using the RMSE and the MAPE (relatively smaller), the AIC, and BIC (Minimum value). Comparing all the models evaluated by ARIMA, ARIMA (1, 1, 1) model satisfies all the coefficient criteria mentioned above. Hence, we consider the model ARIMA (1, 1, 1) is the best model to forecast the future claim's own damage amount in Motor Insurance which is shown in bold form.

TABLE I: Summary results of possible ARIMA Models

Variable Coefficients	RMSE	MAPE	AIC	BIC
ARIMA (1 1 0)	1.6615	15.3128	1.3453	1.3457
ARIMA (0 1 1)	1.4911	13.6273	1.2731	1.2736
*ARIMA (1 1 1)	1.3748	12.0595	1.2599	1.2605
ARIMA (2 1 2)	2.4066	23.1865	1.4992	1.4999
ARIMA (2 1 1)	1.925	18.1192	1.2722	1.2728
ARIMA (1 1 2)	1.5223	13.4652	1.2700	1.2707

*** Best Model**

The mathematical form of generalized ARIMA (1, 1, 1) is represented as

$$X_t = \mu + X_{t-1} + \alpha (X_{t-1} - X_{t-2}) + \beta e_{t-1} \quad (8)$$

Where $\mu = \text{Constant} (1-\alpha)$

X_t, X_{t-1}, X_{t-2} indicates claim period $t, t-1$ and $t-2$ respectively

e_{t-1} indicates residuals of period $t-1$, which constitute a white noise

α, β indicates AR and MA Coefficients of the process respectively

TABLE II: ARIMA (1, 1, 1) Estimates

Parameter	Value	Standard Error	t Statistic	P-Value
Constant	0.0000115	0.00061	0.018	0.985
AR {1}	0.23801	0.00302	78.74	0.000*
Difference	1			
MA {1}	-0.7442	0.00237	-312.76	0.000*

*** Significant**

Thus, the ARIMA (1, 1, 1) model equation (8) can be restated with the estimated values as:

$$X_t = -0.0000115(1 - 0.238) + X_{t-1} + 0.238(X_{t-2} - X_{t-3}) - 0.772e_{t-1} \quad (9)$$

The response variable X_t is dependent on the parameters estimated. Therefore, the equation (9) helps us to estimate future values for prediction of the OD Claim amount. Also, the predicted ARIMA model along with observed OD Claim is shown in Fig 1.

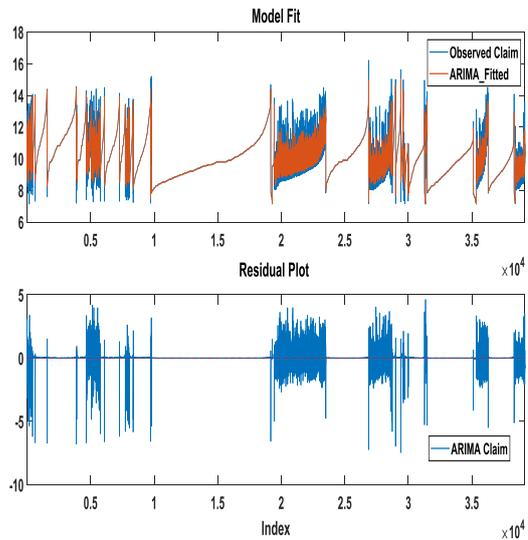


Fig 1: Model fit and Residuals Plot of ARIMA (1 1 1)

In the pursuit of the best model, we evaluated the data by the GLM below.

B. The results of the GLM Model

For the data collected, we have fitted the Generalized linear model (GLM) on the similar lines of ARIMA (1, 1, 1), and various statistics of these models are tabulated in TABLE III.

TABLE III: Summary Results of GLM

Variable Coefficients	RMSE	MAPE	AIC	BIC
GLM	1.183	9.2073	3.174	3.1744

TABLE IV: Model of GLM

	Coefficient	Std Error	t-Statistic	P-Value
β_0	10.514184	0.011956	881.7215	0.000*
β_1	-0.0000269	5	50.95706	0.000*

*** Significant**

Thus, the GLM model equation (1) can be restated as follows:

$$Y_i = 10.5141 - 0.0000269 X_i \quad (10)$$

From the above computations, it is evident that the GLM model fits well as compared to the ARIMA (1 1 1) model concerning the performance characteristics viz; RMSE, MAPE, AIC, and BIC. We know analyze the data by ANN below.

C. The results of the ANN Model

In supervised learning, the output of the network is compared with the desired output (target). For our ANN model, we divided the entire data set into training, testing, and validation.

Data samples of own damage claim, 70% of the data considered as training, 15% of the data has taken for testing, and the remaining 15% is used for validating. Neural Network modeling and estimation are performed using the Neural Network toolbox from MATLAB R2019b software.



The input layers consist of own damage claim amount from different insurance companies. We applied an MLP network consist of multilayer's, that is, one input layer, a single hidden layer, and an output layer.

The neural network function is a two-layer feed-forward network with sigmoid hidden transfer function in the hidden layer and a linear transfer function in the output layer, where each hidden layer consists of 10 neurons, which is configured based on the desired performance. During training, the backpropagation Levenberg-Marquardt algorithm is applied to train the ANN. We have applied TRAINLM, the training performance algorithm and random data division to identify the mean square error (MSE) performance criteria. The number of training epochs was set at 1000, and the control parameter for the algorithm is set to be 0.001. The network is trained until the validation error failed to decrease for six iterations (validation stop). The network performance can be checked from different plots that are used to validate the network model. We retrain the network when performance is not satisfied. Finally, using the TRAINLM method, the training function builds forecast results based on Mean Square Error (MSE) minimization criteria by means of trial and error experiments for own damage claim amount data.

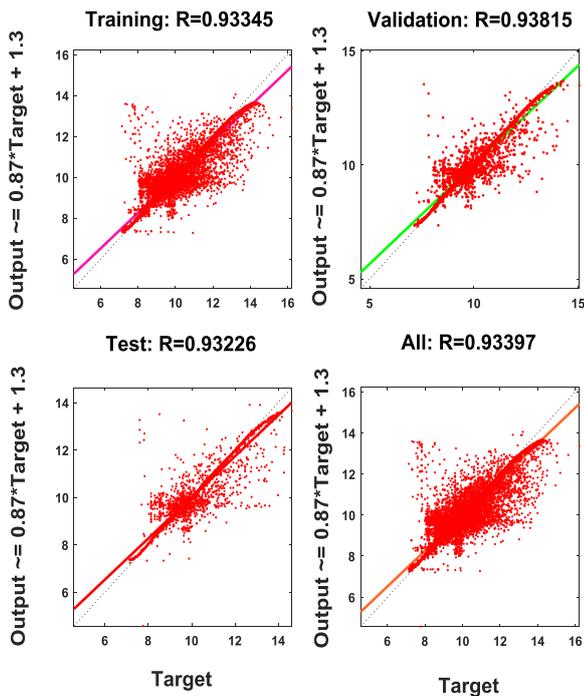


Fig. 2. Regression Plot for ANN model: Training, Test, Validation and Overall data

From Fig. 2, it is observed that Regression plots of ANN-based on training, test, validation, and overall data set to fall with a 45° line with R values in each case as 0.93, which indicates that the predicted plots represent a good fit. It is also indicating that the network predicted outputs are equal to the target. The performance of training validation is shown in Fig. 3, which indicates the best validation performance is 0.17601 at epoch 27.

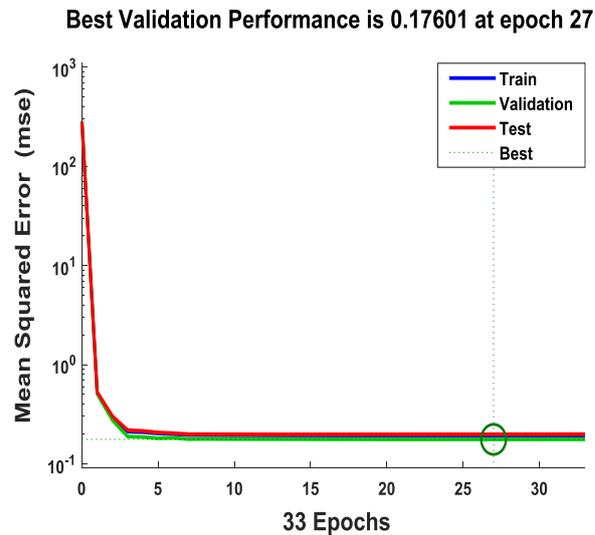


Fig. 3. Performance Plot for ANN model

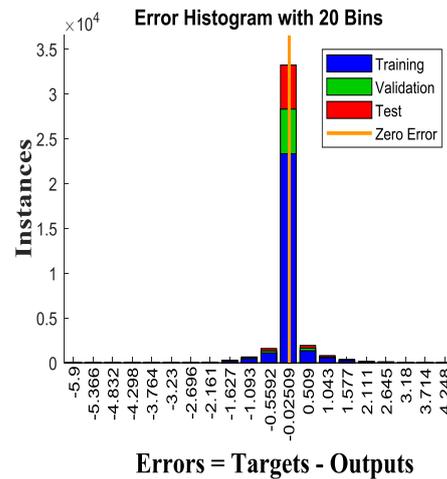


Fig. 4. Error Histogram plot

The error histogram evaluates the network performance of the ANN predicted model and gives an idea about the indication of outliers. From Fig. 4, the training, validation, and testing data are represented by blue, green, and red bars, respectively. In this case, most of the errors fall very close to Zero, which indicates that the prediction was successful with an acceptable margin of error distributions of ANN.

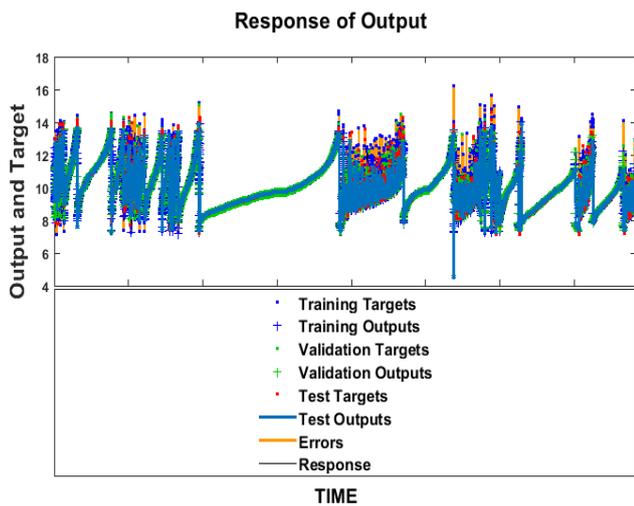


Fig. 5. Response of Predicted target and Error

After identification of the most appropriate model, the fitted model can be used to predict the own damage claim amount for future periods. Fig.5 represents the predicted results of the claim amount with respect to training, validation, and test data set with errors.

D. Comparison of GLM, ARIMA, and ANN

From the above observations, the root means squared error and mean absolute percentage error in testing own damage claim data are used as an evaluation index for predicting the performance of three forecasting models GLM, ARIMA, and ANN. The results of MAPE and RMSE in OD claim amount data calculated using these models are listed in TABLE 5. The model with the lowest MAPE and RMSE was chosen as the appropriate model for forecasting the own damage claim. The exploratory results showed that the ANN model yields a more accurate forecast compared to any other model in each category of vehicles concerning RMSE and MAPE. Hence, the Artificial Neural Network model is the best model for forecasting these TP claim amount.

TABLE V: Comparison of GLM, ARIMA and ANN

Vehicle Categories	Model	Own Damage Claim Data	
		RMSE	MAPE
Own Damage	ARIMA	1.3748	12.0595
	GLM	1.183	9.2073
	ANN*	0.17601	4.2282

* Best Model

V. CONCLUSIONS

This present study is aimed to apply a reliable forecasting model to predict the own damage claim amount in Motor Insurance Segment in Insurance companies of India. Due to the complex nature of data, uncertainty occurrence of the insurance claim, sustained economic growth and ongoing disputes over tariffs, forecasting the own damage claim amount accurately will be beneficial for the insurance

companies for future prospect. In this context, we have developed and compared a time series forecasting techniques such as Autoregressive Integrated Moving Average model, Generalized Linear Model and Artificial Neural Network model based on the performance criteria: RMSE and MAPE. After comparing the above performance criteria along with accurate predictions, the ANN model fits accurately as the right model for the data as compared with other two methodologies with a lesser RMSE and MAPE. This exploratory analysis showed that the ANN model is an effective and appropriate model for predicting the Motor Insurance OD claim data when compared to the GLM and ARIMA.

This comparative data analytics approach would help Insurance companies to have an idea about the expected future claim amounts that can be forecasted with accuracy. Also, this modeling approach and forecasts will help the Insurance Companies to Budget their future revenue planning for the ensuing years for an easy disbursement of the Claims.

REFERENCES

- Vipul k. Dabhi and Sanjay Chaudhary (2016). "Financial Time Series Modeling and Prediction Using Postfix-GP". Computational Economics, Volume 47, Issue 2, PP: 219-253 DOI:10.1007/s10614-015-9482-y. <https://ideas.repec.org/a/kap/compec/v47y2016i2p219-253.html>
- J. David Cummins and Gary L. Griepentrog (1985). "Forecasting Automobile Insurance Paid Claim Cost", Elsevier Science Publishers, B V (North-Holland). <https://www.sciencedirect.com/science/article/pii/0169207085900032>
- G.C.Mesike, I. A. Adeleke and Ade Ibiwoye (2012). "Predictive Actuarial Modeling of Health Insurance Claims Costs", International Journal of Mathematics and Computation, ISSN: 0974-5718, Vol.14, issue No 1. <http://www.ceser.in/ceserp/index.php/ijmc/article/view/2545>
- Cvetko Andreeski, Bratislav Milosevic, Vladimir Njegomir (2012), "Analysis of the Life Insurance Market in the Republic of Macedonia", Economic Annals, Volume LVII, ISSN: 0013-3264. <http://www.ekof.bg.ac.rs/wp-content/uploads/2014/06/215>
- Soliman Osama R (2003). "Forecasting Loss Ratio in Property and Liability Insurance companies using Autoregressive and Integrated Moving Average Models (ARIMA) for Time-series Analysis", New Horizons and Journal, El monophobia University.
- Tarek Taha (2017). "Forecasting Fire Insurance Loss Ratio in Misr Insurance Company", El-Bahith Review, Issue 17 (17); 31-39. https://scholar.google.co.in/scholar?hl=en&as_sdt=0%2C5&as_vis=1&q=New+Horizons+and+Journal+%2C+El+monophobia+University&btnG=
- Samuel Erasmus Alnaa and Ferdinand Ahiakpor (2011). "ARIMA (autoregressive integrated moving average) approach to predicting inflation in Ghana", Journal of Economics and International Finance, vol.3(5), ISSN 2006-9812. https://www.academia.edu/30842596/ARIMA_autoregressive_integrated_moving_average_approach_to_predicting_inflation_in_Ghana
- Thabani Nyoni (2019). "Modeling and Forecasting Inflation in Burundi using ARIMA Models", Munich Personal Repec Archive, MPRA Paper No. 9244, 2019. <https://mpra.ub.uni-muenchen.de/92444/>
- Ina Khandelwal, Ratnadip Adhikari and Ghanshyam Verma (2015). "Time Series Forecasting Using Hybrid ARIMA and ANN Models based on DWT Decomposition", Science Direct, Procedia Computer Science 48,173-179. <https://cyberleninka.org/article/n/391592.pdf>
- Olayiwola Olaniyi Matthew, ApantakuFadeka Sola, Bisira Hammed Oladiran & Adewara Adedayo Amos (2013). "Prediction of Stock Price using Autoregressive Intergrated Moving Average Filter (ARIMA(P D Q))", Global Journal of Science Frontier Research Mathematics and Decision Sciences, Volume 13, Issue 8, ISSN: 0975-5896.



<https://pdfs.semanticscholar.org/6940/a88b1199371c340fe1be3f201802fc6d26da.pdf>

11. Carina Intan Permatasari, Wahyudi Sutopo and Muh. Hisjam (2018). "Sales Forecasting Newspaper with ARIMA: A Case Study", AIP Conference Proceedings 1931, 030017. <https://aip.scitation.org/doi/10.1063/1.5024076>
12. Z.Asha Farhath, B. Arputhamary and Dr. L. Arockiam (2016). "A Survey on ARIMA Forecasting Using Time Model, International Journal of Computer Science and Mobile Computing", Vol.5, Iss.8, ISSN:2320-088X. <https://www.ijcsmc.com/docs/papers/August2016/V5I8201626.pdf>
13. Jamal Fattah, Latifa Ezzine, Zineb Aman, Haj El Mousami and Abdeslam Lachhab(2018). "Forecasting of demand using ARIMA Model", International Journal of Engineering Business Management, Vol 10: 1-9, 2018. <https://journals.sagepub.com/doi/full/10.1177/1847979018808673>
14. Ch Madhavi Latha, K Siva Nageswararo, M Venkataramanaiah (2018). "Forecasting Time Series stock returns using ARIMA: Evidence from S& P BSE Sensex", International Journal of Pure and Applied Mathematics, Volume 118, No. 2, ISSN: 1314-3395. <https://acadpubl.eu/hub/2018-118-24/2/223.pdf>
15. J C Ramesh Reddy, T Ganesh, M Venkateswaran, PRS Reddy (2017). "Forecasting of Monthly Mean Rainfall in Rayalaseema", International Journal of Current Research and Review, Vol 9, Issue 24, 2017. http://ijcrr.com/uploads/2403_pdf.pdf
16. Mohamed Reda Abonazel and Ahmed Ibrahim Abd Elftah (2019). "Forecasting Egyptian GDP using ARIMA Models", Reports on Economics and Finance, Vol.5, no.1, PP: 35-47. <https://doi.org/10.12988/ref.2019.81023>
17. Choden and Suntaree unhapipat (2018). "ARIMA model to forecast international tourist visit in Bumthang, Bhutan", Journal of Physics; conference series (1039) 012023, 2018. <https://iopscience.iop.org/article/10.1088/1742-6596/1039/1/012023>
18. Mohammed H. Alsharif, Mohammad K. Younes and Jeong Kim (2019). "Time Series for Prediction of Daily and Monthly Average Global Solar Radiation: The Case Study of Seoul, South Korea", Journal of symmetry, 11, 240; doi: 10.3390/sym 11020240. <https://www.mdpi.com/2073-8994/11/2/240>
19. S. Raid, J. Mania, L.Bouchaou and Y. Najjar (2004), "Rainfall-runoff model using an artificial neural network approach", Mathematical and Computer modeling: An International Journal, Vol.40, No. 7-8, October 2004.
20. Suppachai Pathumnakul, Piewthongngam and Arthit Apichottanakul (2009), "A neural network approach to the selection of feed mix in the feed industry", Computers and Electronics in Agriculture, Vol. 68, Issue 1, P.No. 18-24, August 2009.
21. Arthit Apichottanakul, Suppachai Pathumnakul and Piewthongngam(2009), "Using an artificial neural network to forecast the market share if Thai Rice", IEEE-International Conference on Industrial Engineering and Engineering Management, ISBN:978-1-4244-4899-2, 2010. DOI: 10.1109/IEEM.2009.5373247.
22. Ying Z and Hanbin X (2010), "Study on the model of demand forecasting based on artificial neural network" Distributed Computing and Applications to Business Engineering and Science (DCABES), 9th International Symposium, IEEE, pp. 382386, 2010.
23. Karin Kandananond (2012), "A comparison of various forecasting methods for Autocorrelated Time Series", International Journal of Engineering Business Management, Vol. 4, 2012. DOI: 10.5772/51088.
24. Kandananond(2011), "Forecasting Electricity Demand in Thailand with an Artificial Neural Network Approach", Energies, ISSN: 1996-107, Vol.4, 2011. Doi: 10.3390/en4081246
25. Nadia Slimani, Ilham Slimani, Nawal ISbiti and Mustapha Amghar (2019), "Traffic forecasting in Morocco Using Artificial Neural Networks", Procedia Computer Science, Volume 151, Pages 471-476, 2019. <https://doi.org/10.1016/j.procs.2019.04.064>
26. E.T. Lau, L. Sun and Q Yang (2019), "Modelling, prediction and classification of student academic performance using artificial neural networks", SN Applied Sciences, Springer Nature, 1:982, 2019.
27. Box, George E.P and Gwilym M Jenkins., (1976). "Time Series Analysis: Forecasting and Control", Revised Edition, Oakland , CA: Holden-Day. <https://trove.nla.gov.au/work/10907520?q&online=true>.
28. Vandaele .W (1983). "Applied time Series and Box- Jenkins models", New York, Academic Press,Orlando: London. <https://trove.nla.gov.au/ version/45004313>.

29. Buckman Akuffo & Enoch Mintah Ampawz(2013). An Autoregressive Integrated Moving Average (ARIMA) Model for ghana's inflation (1985 – 2011) Mathematical Theory and Modeling, ISSN 2224- 5804, VolumeNo:3. <https://iiste.org/Journals/index.php/MTM/article/viewFile/4909/4987>
30. T W Anderson (1971). The Statistical Analysis of Time Series, Wiley Classic Library Edition <https://trove.nla.gov.au/version/261201497>.

AUTHORS PROFILE



Mr. V SelvaKumar, Research Scholar in Statistics in the area of Data Analytics, Department of Mathematics, Birla Institute of Technology & Science- Pilani, Hyderabad Campus. He is working as an Assistant Professor in the Department of Mathematics and Statistics, Bhavan's Vivekananda College of Science, Humanities & Commerce. To his credit he has published 10 papers in different national and International journals. Also, presented 6 papers in national and International conferences. His areas of interest are Data analytics, Time Series Analysis and Design of Experiments.



Prof. Dipak K Satpathi, is Associate Professor, Head, Department of Mathematics, Birla Institute of Technology & Science, Pilani, Hyderabad Campus. He received his B.Sc. and M.Sc. from University of Burdwan, India. He received his Ph.D. degree in Mathematics from IIT Kanpur, India. His areas of interest are Mathematical Modeling of Biological Systems, Fluid Mechanics and Financial Mathematics. He has authored a book "Mathematical Modeling: Application, Issues and Analysis" (Published by Ane Books India, New Delhi, India) in the year 2007. He has published 15 papers and presented 13 papers in various national and international conferences. He guided one Student for PhD.



Dr. P.T.V Praveen Kumar, is Assistant Professor in the Department of Mathematics, Birla Institute of Technology & Science, Pilani, Hyderabad Campus. He obtained his bachelor's and Master's degrees from Sri Krishnadevaraya University. He received his Ph.D. degree in Statistics from the Sri Krishnadevaraya University, India. His areas of interest are Multivariate Data Analysis and Bio statistics. He has published 7 papers in the area of Data Analytics and Bio statistics. Also, he has presented papers in both international and national conferences.



Prof. V V Hara Gopal, Visiting Professor, Dept. of Mathematics, Birla Institute of Technology & Science, Hyderabad Campus. He was a Former Director, Centre for Quantitative Methods, Osmania University, Hyderabad. He has done his M. Sc. (Statistics), PhD from Osmania University, Hyderabad, India. He has supervised twenty-Five PhD dissertations and seven MPhil dissertations in the areas of Statistics, Operations Research, Image Compression, Group testing, Econometrics, Textual Mining and Social Statistics. He published around 110 research papers in national and International Journals. He has successfully completed many government projects from different states in India. He is life Member in many Scientific Institutions. He is member of Editorial Board-International Journal of Agricultural and Statistical Sciences, and Editor for Telugu Academy books.