DOI: 10.1111/rmir.12200

FEATURE ARTICLE

Risk Management and Insurance Review WILEY

Using an adaptive network-based fuzzy inference system model to predict the loss ratio of petroleum insurance in Egypt

Ahmed A. Khalil[1,2] | **Zaiming Liu**[1] | **Attia A. Ali**[2]

[1]School of Mathematics and Statistics, Central South University, Changsha, Hunan, China

[2]Faculty of Commerce, Assiut University, Assiut, Egypt

Correspondence
Ahmed A. Khalil, Faculty of Commerce, Assiut university, Assiut 71515, Egypt.
Email: ahmedrohim91@aun.edu.eg

Abstract

Insurance companies and those interested in developing insurance services seek to use modern mathematical and statistical methods to study further and analyze all the company's corporate internal and external performance indicators. Loss ratio is a vital indicator used to measure performance and predict future losses in insurance companies. Many pivotal processors, such as underwriting and pricing depending on it. Therefore, accurate predictions assist insurance companies in making decisions properly. Thus, this paper aims to use the adaptive network-based fuzzy inference system (ANFIS) and autoregressive integrated moving average (ARIMA) models in forecasting the loss ratio of petroleum insurance in Misr Insurance Holding Company from 1995 to 2019. We applied many ANFIS models according to ANFIS properties and used the first 21 years (1995–2015), making up the training data set, which represents 85% of the data, as well as the past 4 years (2016–2019). Which are used for the testing stage and represent 15% of the data. Our finding concluded that ANFIS models give more accurate results than ARIMA models in predicting the loss ratio during the investigation by comparing results using predictive accuracy measures.

KEYWORDS

ANFIS, ARIMA, Egypt, forecasting, loss ratio, neuro-fuzzy system

1 | INTRODUCTION

The core of insurance companies works toward achieving the goal of providing insurance protection to the best of their ability in a way that satisfies customers (Kasturi, 2006; Safwat, 2003). Apart from their contribution to achieving a company's profit, maintaining financial position integrity, and improving competitive position in the market, insurance companies must concern themselves with the competition. That is particularly true in Egypt, where companies must contend with competition on domestic and foreign fronts while also dealing with recent economic changes in the Egyptian insurance market (Cappiello, 2020; Ibrahim et al., 2011).

The loss ratio (LR) is vital to measure the efficiency of the insurance company's functions, namely, underwriting, pricing, reinsurance, investment, and claims compromise. The insurance sector faces significant volatility, described according to the inherent risks it covers, along with the significant losses in the covered risks (El-Bassiouni, 1991). The accuracy of an LR estimate helps general insurance companies avoid possible losses if the risk occurs. Therefore, our study is concerned with the importance of accurate LR estimation.

The autoregressive integrated moving average model (ARIMA) is one of the conventional models that have been used and accepted in many insurance studies that have been applied in Egypt. Most of these studies have concluded that the ARIMA model is a common traditional statistical method that can be used for acceptable forecasting results, with univariate time series data compared to other conventional statistical models (Hamid & Mohamed, 2015; Safwat, 2003; Soliman, 2003).

Nowadays, many human knowledge fields have turned to modern artificial intelligence techniques and are applied in many scenarios rather than relying on conventional techniques, such as ARIMA for prediction. This is due to two factors. First, conventional techniques have some drawbacks, such as the fact that most conventional techniques rely on certain assumptions about variables. An example of this is the assumption of linearity between dependent and independent variables. Thus, if the relationship between the independent and the dependent variables is not known or the assumptions are not fulfilled, the model becomes inappropriate for the prediction process and more sensitive to outliers (Akkoç, 2012). The second factor is related to insurance data characteristics, such as uncertainty, noisiness, and incomplete information (Yunos et al., 2019). Therefore, it is essential to use modern prediction methods, such as the neuro-fuzzy system (NFS) models. These models aim to combine the advantages of both the neural network (NN) and fuzzy logic (FL) approaches collectively and employ logic in their operations instead of ongoing relationships between variables. One of the popular NFS prediction models is the adaptive network-based fuzzy inference system (ANFIS) that we will be using in this study.

ANFIS is utilized in a few investigations in various fields. An example of this is Çakıt et al. (2020), who utilized ANFIS to model the Japanese petrochemical industry's safety culture. The use of ANFIS indicated that employee attitudes were the most critical predictor of behaviors characteristic of code violations and personnel errors, while coworker support was the most fundamental indicator in motivation modeling related to personnel safety. Oroian (2015) assessed the impact of temperature dampness and the recurrence on nine viscoelastic honey samples by using artificial neural networks (ANN) and ANFIS and was able to demonstrate that the ANNs methodology is a superior predictor of the evolution of viscoelastic factors of honey in the function of high temperature, frequency, and wetness content than ANFIS. Akkoç (2012) proposed a three-stage hybrid ANFIS for credit score analysis and demonstrated that the suggested model reliably implements superior to the logistic regression analysis, linear discriminant analysis, and ANN methods. Abbasi et al. (2014) attempted to make a model that employs minimal data measures to predict the average monthly discharge in the Jajrud River using ANFIS models. The consequences of this investigation show the ability of ANFIS to predict river runoff. However, in the insurance field, there is little research in applying this. Therefore, our study aims to apply the ANFIS and ARIMA models so that they can be used to predict the petroleum insurance's LR in the Misr Insurance Holding Company, while also comparing their performance using three quantitative standards for analytical performance and evaluation measures. Mean-squared error (MSE), root-mean-squared error (RMSE), and mean absolute error (MAE) are used to validate all models. The LR data are collected from the Egyptian Financial Supervisory Authority's annual report for insurance companies from 1995 to 2019.

2 | THEORETICAL FRAMEWORK

This section is divided into two subsections: Section 2.1 defines LR, its importance, and its uses. Section 2.2 explains the NFS term and outlines the advantages and disadvantages of the NFS components: NN and FL.

2.1 | LR

The National Association of Insurance Commissioners' annual statement defines LR as "A proportion of the relationship between claims and premiums." While this definition is straightforward, we should inspect the components that impact both the numerator and denominator of this proportion. LR is the proportion of losses incurred in claims divided by premiums earned. This ratio is one of the most critical performance factors for insurance companies as LR shows the effectiveness of an insurance company's performance. In this study, the LR is calculated by dividing incurred claims by earned premiums. Thus, insurance companies that consistently experience high LR may be in dire financial health. That is an indication that they are not collecting enough premiums to pay claims. In general, all insurance companies propose is that their premiums increase and that the claims that they are required to compensate decrease (Berhe & Kaur, 2017).

The LR can be calculated from the following formula:

$$\text{Loss ratio(LR)} = \frac{\text{Loss incurred}}{\text{Premiums earned}}.$$

The extent of the movement of these variables in the future depends on their experience. In most cases, the experience is not equal to what would be expected of them based on historical data (Hogg & Klugman, 2009).

2.1.1 | Uses of the LR

There are many uses for LRs. The LR can be used as a relative cost indicator, with the ratio reflecting the percentage of premiums returned to customers in the form of payments or compensation. This shows the direct benefits that insurance customers receive in exchange for installment payments (Abd El-Zahir, 2015). Further, LRs can be used as a measure of profitability, as profit ratios correlate inversely with LRs. This means that a higher LR would lead to a decrease in income underwriting and vice versa. The LR is also used instead of the profit ratio because the LR is a source of uncertainty when determining insurance profits (Lamm-Tennant et al., 1992).

Moreover, the average LR is used as an indicator of underwriting risk related to underwriting results' uncertainty or unpredictability. The fluctuation in the LR is almost variable, reflecting the possibility based on the results achieved. The insurance branch faces underwriting risk if the variation in the LR is significant. Accordingly, the temporary change in the LR is used as a measure of the total underwriting risk.

2.1.2 | Importance of predicting the LR for insurance companies

The importance of predicting the LR stems from being one of the criteria for measuring insurance companies' performance and is used to set goals and establish policies related to all activities associated with the production, reinsurance, and technical allocations. Predicting the future loss rate is not as easy as it is inaccurate due to random errors. As a result of many other factors, such as its dependence on the personal assessment factor and reliance on previous data, a carefully studied scientific method predicting the expected LR helps insurance companies. This approach addresses the effects of inflation and measures underwriting profitability, facilitating the accurate forecasting of underwriting profits in the future. Further, it allows for the prediction of future written premiums and contributes to the stability of LRs in the long term through accurate and continuous forecasting in the short term. It also determines the rate of surplus insurance activity for branches inside the insurance companies. Further still, this approach enables insurance companies to make decisions related to underwriting, pricing, and reinsurance while also rationalizing the decision-making process, including planning the underwriting profit margin and setting the underwriting policy accordingly forecasted ratios (Bakhit et al., 2004).

2.2 | NFS

Incorporating the NN and FL has created a new research field called fuzzy neural system. This field began at the end of the 1980s and has contributed to significant growth while being widely used in different approaches. The term "neural fuzzy" is a kind of system characterized by its identical fuzzy controller structure, where the fuzzy sets and rules are amended using NNs that repeatedly tune methods using data vectors (input and output system data) (Vieira et al., 2004). A learning machine, another important term, finds a fuzzy system's parameters using approximation methods from NN (Kruse, 2008).

The fuzzy neural system aims to combine the benefits of both the NN and FL approaches collectively, where each AI technique has an especial ability. ANN precedes machine learning by emulating a neural system of humans. FL quite closely mimics a human's thinking. However, these techniques have some specific drawbacks, in the case of ANN, it is difficult to interpret the result obtained, while for FL doesn't have the ability to learn. Therefore, these techniques' drawbacks are eliminated when integrated into one model capable of gathering the advantage of learning artificial neural networks and the modeling supremacy of FL (Akkoç, 2012; Vieira et al., 2004). The most common architecture used for fuzzy neural systems uses NNs to learn a stable structure (Nauck et al., 1997). ANFIS belongs to the combined NFS, which was presented by Jang (1993).

Thus, the advantage of a fuzzy neural approach can be summarized by the following points:

- All the parameters of the function between inputs and outputs are usually not fully known.
- In contrast to simple neural or fuzzy techniques, the fuzzy neural system technique has merged advantages of neural and fuzzy techniques.

This study aims to investigate the ability of the ANFIS model to predict LR with accuracy. Further, the proposed model's performance will compare with the ARIMA model.

3 | DATA AND METHODOLOGY

This section describes LR data, which we apply in our study and use to clarify the models that we are utilizing to forecast LRs. We also cover the accuracy measures that we will use to compare the models.

3.1 | Data

This study covers 25 years from 1995 to 2019, depending on the annual data of the LR of the most significant national insurance company (Misr Insurance Holding Company). Misr Insurance Holding Company acquires more than 95% of petroleum insurance within the Egyptian insurance market, according to the most recent report issued by the Egyptian Financial Supervision Authority (EFSA). This data of the proposed work is publicly available online on the Egyptian Financial Supervisory Authority's annual report for insurance companies.[1] Table 1 and Figure 1 present the data of the LR of the petroleum insurance branch (input) within the general insurance sector during the period of study from 1995 to 2019.

The descriptive statistics related to the data of the LR of petroleum insurance are presented in Table 2.

3.2 | Methodology

This study attempts to construct a model that depends on the ANFIS model to predict the LR of the Misr Insurance Holding Company's petroleum insurance branch.

3.2.1 | ANFIS

The architecture of the ANFIS is one of the first NFS models introduced by Jang (1993) and implements the Takagi-Sugeno inference system. Figure 2 shows the ANFIS architecture, which consists of five layers. The Sugeno system has many advantages, such as being computationally effective and suitable to work with linear, optimization, and adaptive methods. We

[1]https://fra.gov.eg/%d8%a7%d9%84%d8%aa%d9%82%d8%a7%d8%b1%d9%8a%d8%b1-%d8%a7%d9%84%d8%b3%d9%86%d9%88%d9%8a%d8%a9/?doing_wp_cron=1635874193.5824239253997802734375

TABLE 1 Loss ratio of the petroleum insurance branch

Year	1995	1996	1997	1998	1999	2000	2001	2002	2003
LR (%)	5.58	3.56	2.83	0.46	6.44	69.99	6.19	11.83	24.93
Year	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**
LR (%)	49	427.4	12.08	15.74	−49.03	−25.6	23.03	25.59	51.82
Year	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**		
LR (%)	55.2	48.21	33.92	72.54	114.72	56.84	80.25		

FIGURE 1 Loss ratio of the petroleum insurance branch

TABLE 2 Descriptive statistics of loss ratio for the petroleum insurance branch

Descriptive statistics	
N	25
Minimum	49.0
Maximum	427.47
Mean	44.94
Standard deviation	87.17

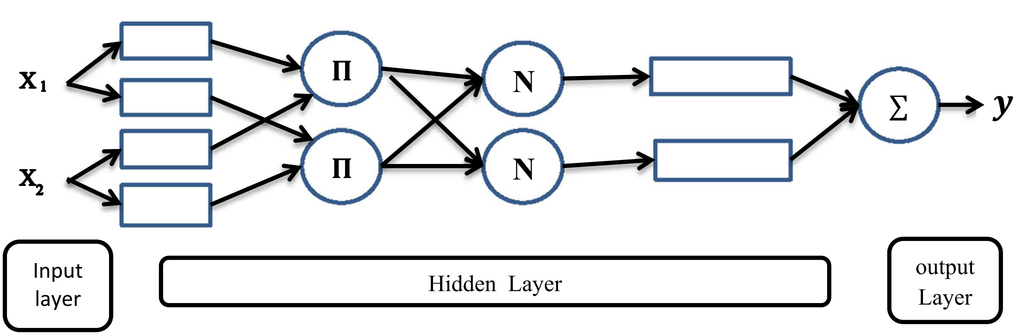

FIGURE 2 The architecture of the adaptive network-based fuzzy inference system

assume, for example, a FIS with two inputs x_1 and x_2 and y output. Figure 3 shows the first-order Sugeno fuzzy model; we can represent a standard rule set with two "If–then" fuzzy rules as follows (Akkoç, 2012; Jang, 1993; Jovanovic et al., 2004; Oroian, 2015):

$$\text{Rule } 1: \text{ If } x_1 = A_1 \text{ and } x_2 = C_1 \text{ then } f_1 = n_1x_1 + m_1x_2 + r_1,$$

$$\text{Rule } 2: \text{ If } x_1 = A_2 \text{ and } x_2 = C_2 \text{ then } f_2 = n_2x_1 + m_2x_2 + r_2,$$

where x_1 and x_2 are independent variables, A_i and C_i are fuzzy sets, and n_i, m_i, and r_i are dependent variable parameters.

Layer 1: This layer is accountable for the mapping of the input variable relative to each membership function. Every one of g nodes in this layer is an adaptive node with a node function described by:

$$O_g^1 = \mu A_g(x), \quad \text{for} \quad g = 1, 2, ... \tag{1}$$

where x is the input node, A_g is the linguistic labels (small, medium, and large) linked by this node function, and O_g^1 is the membership functions of A_g.

Layer 2: All nodes in this layer are fixed nodes labeled **Π**, which doubles the coming signals and sends the product out. The outputs of this layer, which represents the firing strength of the rules, can be denoted as:

$$O_{2,g} = w_g = A_g(x_1) \times C_g(x_2), \quad \text{for} \quad g = 1, 2, ... \tag{2}$$

Layer 3: All nodes in this layer are fixed nodes. Each node is labeled N. The ith node computes the ith rule firing strength to the sum of all rules firing strength.

$$O_{3,g} = \overline{w_g} = \frac{w_g}{w_1 \quad + \quad w_2}, \quad \text{for} \quad g = 1, 2 \tag{3}$$

Layer 4: The consequence is computed in this layer, where each node estimates the contribution of th rule about the overall output:

$$O_{3,g} = \overline{w_g} f_g = = \overline{w}(n_g x_1 + m_g x_2 + r_g) \tag{4}$$

Where $\overline{w_g}$ is the output of layer three and n_i, m_i, and r_i make up the parameter set.

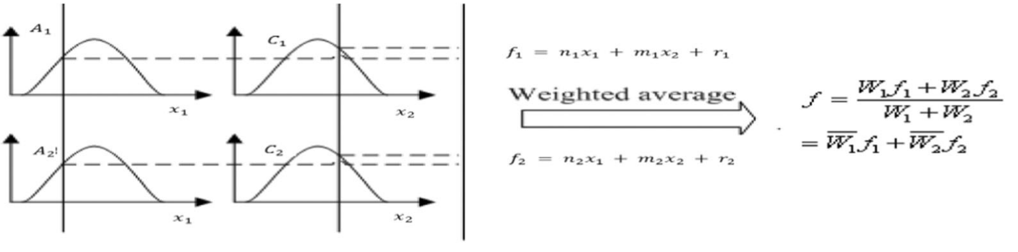

FIGURE 3 Two inputs for the first-order Sugeno fuzzy model with two rules

Layer 5: In this layer, the single node is a fixed node labeled that calculates the overall output as the summation of all coming signals:

$$O_{5,g} = \sum_{1} \overline{w_g} f_g = \frac{\sum_1 w_g f_g}{\sum_1 w_g}. \tag{5}$$

3.2.2 | ARIMA

ARIMA model is one of the famous linear statistical models for time series forecasting. The ARIMA model is used in time arrangement estimations and parameter estimations and has led to many other popular techniques (Box & Jenkins, 1976) (Table 3).

There are three parameters in the ARIMA (p, d, and q) model. Parameter p is related to AR (p), parameter d is related to I (d), and parameter q is related to MA (q). This model depends on past data and is used to make predictions. The ARIMA model can interpret for time-based dependence in some ways (Geetha & Nasira, 2016).

- First, by taking d differences, the time series will make it stationary. If $d = 0$, the observations are straight modeled, and if $d = 1$, the variations between the sequence observations are modeled.

- Second, the stationary process time dependence is modeled by the p autoregressive model. The equation for p shall be:

$$X_t = a + \sum_{i=1}^{\mathrm{p}} \phi_i X_{t-1} + \varepsilon_t, \tag{6}$$

  where a is the constant, ϕ_i is the model parameter, X_t is the observed value at t, and ε_t is the random error.
- Third, q is the average moving term and all-time variations. It indicates the observation of previous errors. The formula for q is:

$$X_t = \varepsilon_t + \sum_{i=0}^{\mathrm{q}} \phi_j \varepsilon_{t-1}, \tag{7}$$

  where ϕ_i is parameter model, and e_t is an error term.
- We get the ARIMA model from these three models. The standard type of ARIMA models is defined by the following:

$$y_t = a_0 + \sum_{i=1}^{\mathrm{p}} \phi_i\, y_{t-1} + \sum_{j=0}^{\mathrm{q}} \phi_j \varepsilon_{t-j}, \tag{8}$$

TABLE 3 Autoregressive integrated moving average (ARIMA) models

ARIMA models	MSE	RMSE	MAE
ARIMA (1,0,1)	0.8203	0.9057	0.443
ARIMA (1,0,2)	0.7126	0.8442	0.448
ARIMA (2,0,1)	0.6988	0.8359	0.442
ARIMA (1,1,1)	0.9279	0.9633	0.475
ARIMA (2,1,1)	0.9270	0.9628	0.471

Abbreviations: MAE, mean absolute error; MSE, mean-squared error; RMSE, root-mean-squared error.

where y_t stationary is a process of stochastic, a_0 is a constant, ε_t represent the error, ϕ_j represents the coefficient of autoregression, and ϕ_j is the average moving coefficient.

3.2.3 | Accuracy of forecasting measures

To evaluate the forecast data accuracy, some error measures are used to evaluate the forecast procedure. Three of the widespread errors that are used to evaluate the accuracy are MSE, RMSE, and MAE. All of these measures can be computed by using the following equations (Abbasi et al., 2014; Oroian, 2015):

MSE is the average of the error squares. The error differs from the predicted value to the actual value and is defined as:

$$\mathrm{MSE} = \frac{1}{n}\sum_{t=1}^{n}\left(y_t - y_t^*\right)^2. \tag{9}$$

RMSE is a standard error-index statistics used to determine the difference between the predicted model values and those of the model observed (Lin et al., 2006; Nayak et al., 2004) and is defined as:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(y_t - y_t^*\right)^2}. \tag{10}$$

MAE is measured utilizing a term-by-term comparison of the relative error in the variable's actual prediction and defined as:

$$\mathrm{MAE} = \frac{1}{n}\sum_{t=1}^{n}\left|y_t - y_t^*\right|, \tag{11}$$

where y_t and y_t^* indicate samples of current and predicted model data, respectively. The sample size is n.

4 | RESULTS AND DISCUSSION

This section provides implementation details for the proposed models, which provide a better result for forecasting the LR of the petroleum insurance branch in Misr Insurance Holding Company.

4.1 | ARIMA models

For the simulation of ARIMA, IBM SPSS Statistics 23 was used. SPSS is a complete and useful statistical analysis tool that provides the best ARIMA model for observed LR data. According to a plot autocorrelation function (ACF) and a partial autocorrelation (PACF) for the LR data, we applied many ARIMA models. We get moving average terms from the ACF, and from the PACF, autoregressive terms. The ARIMA (2, 0, 1) is the best fit model using MSE = 0.6988, RMSE = 0.8359, and MAE = 0.442, as our values. The best fit ARIMA model is presented in Figure 4.

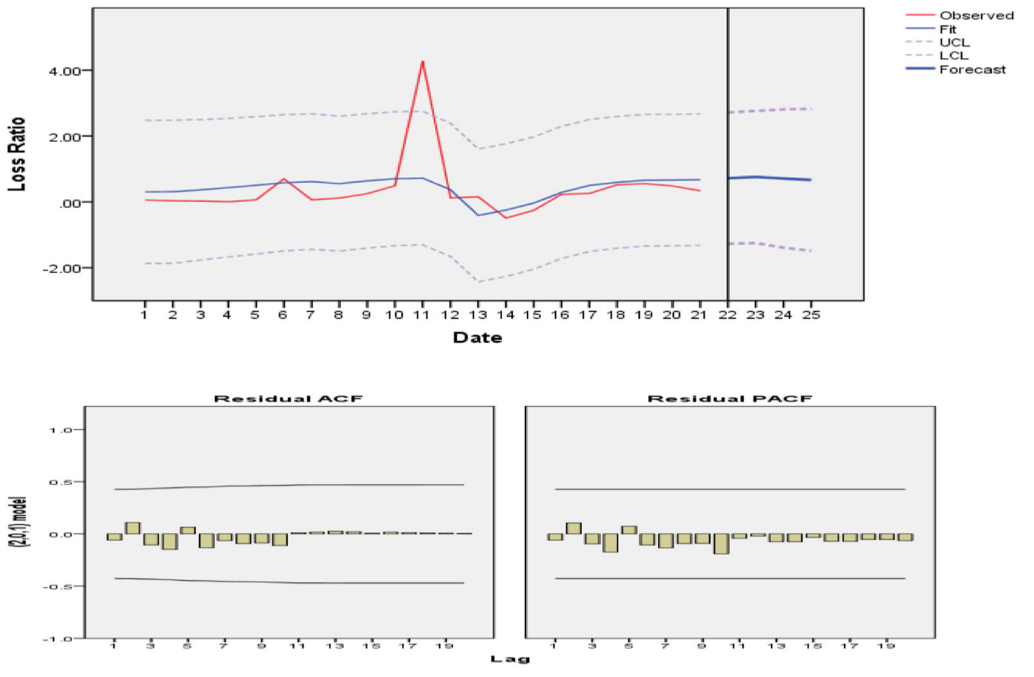

FIGURE 4 Autocorrelation function (ACF) and plot ACF (PACF) for autoregressive integrated moving average (2, 0, 1) model residuals

4.2 | ANFIS model

MATLAB fuzzy designer is used for ANFIS modeling. We applied many ANFIS models according to ANFIS properties, such as membership function (MF)-type inputs or outputs, and optimal train FIS methods using the observed data. The first 21 values (1995–2015) constituted the training data set, while the last 4 years (2016–2019) were used for the testing stage. Grid fuzzy partition was used to establish the rule-based relationship between the input and output variables. The number of membership functions for each input of ANFIS was set to three. The MF types for the inputs selected are a Gaussian curve, a trapezoidal-shaped function, and a triangular-shaped function. These functions utilize an optimal train, FIS method hybrid, and back-propagation, respectively.

The following discussion details the results obtained. Time data and the actual value were used as an input and an output, respectively, with three membership functions for each input as a default for the layer (2,3,4) (1-3-3-3-1 architecture). The best fit ANFIS model was chosen using MSE, RMSE, and MAE. The results showed that model number five of the ANFIS models were the best model, by using each a Gaussian membership function for its inputs, linear for its outputs, and a hybrid method for train FIS. As shown in Figure 5, the variation of error of model (5) with increasing epoch numbers, it found that errors continue to diminish until approximately 504 epochs, before becoming almost constant. The results of the ANFIS models and their specifications have been summarized in Table 4.

Finally, Table 5 shows ANFIS and ARIMA models' performance comparisons due to MSE, RMSE, and MAE in training and test data. We can see that the ANFIS model is lower than the ARIMA model in MSE, RMSE, and MAE. Consequently, the above findings show that ANFIS works better than ARIMA, as shown in Figure 6 and Table 6.

FIGURE 5 Variation of error of the best adaptive network-based fuzzy inference system model

TABLE 4 Adaptive network-based fuzzy inference system (ANFIS) models

	MF type		Train FIS		MSE		RMSE		MAE	
ANFIS models	Input	Output	Optimal. method	Epochs	Train	Test	Train	Test	Train	Test
1	Trimf	Linear	Back-prop	4000	0.043	0.477	0.208	0.690	0.151	0.524
2	Gaussmf	Linear	Back-prop	4000	0.067	0.422	0.259	0.649	0.180	0.614
3	Trapmf	Linear	Back-prop	4000	0.050	0.479	0.225	0.692	0.148	0.660
4	Trimf	Linear	Hybrid	4000	0.549	0.944	0.741	0.972	0.469	0.832
5	Gaussmf	Linear	Hybrid	4000	0.040	0.067	0.201	0.260	0.111	0.229
6	Trapmf	Linear	Hybrid	4000	0.580	0.297	0.761	0.545	0.426	0.503

Abbreviations: FIS, fuzzy inference system; MAE, mean absolute error; MF, membership function; MSE, mean-squared error; RMSE, root-mean-squared error.

TABLE 5 ANFIS and ARIMA models' performance comparisons due to MSE, RMSE, and MAE

Measure	MSE		RMSE		MAE	
Model	Training	Testing	Training	Testing	Training	Testing
ANFIS	0.040	0.067	0.201	0.260	0.111	0.229
ARIMA	0.698	0.069	0.835	0.264	0.442	0.239

Abbreviations: ANFIS, adaptive network-based fuzzy inference system; ARIMA, autoregressive integrated moving average; MAE, mean absolute error; MSE, mean-squared error; RMSE, root-mean-squared error.

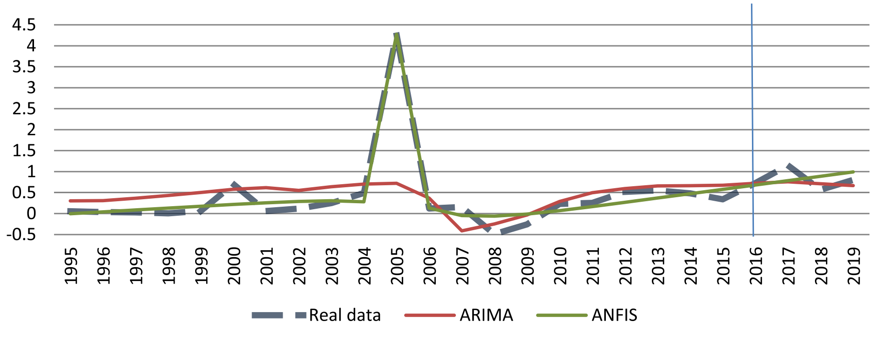

FIGURE 6 Comparison of autoregressive integrated moving average (ARIMA) and adaptive network-based fuzzy inference system (ANFIS) model

TABLE 6 Comparison of actual value and predicted value between ARIMA and ANFIS models

		Predicted value		Residual	
Year	**Actual value**	**ANFIS model**	**ARIMA model**	**ANFIS model**	**ARIMA model**
1995	0.0558	−0.003449989	0.3037	0.0592	−0.2479
1996	0.0356	0.041172992	0.3092	−0.0056	−0.2736
1997	0.0283	0.085750365	0.3668	−0.0575	−0.3385
1998	0.0046	0.130196441	0.4326	−0.1256	−0.4280
1999	0.0644	0.174271246	0.5036	−0.1099	−0.4392
2000	0.6999	0.21731629	0.5809	0.4826	0.1190
2001	0.0619	0.257569785	0.6169	−0.1957	−0.5550
2002	0.1183	0.290497665	0.5535	−0.1722	−0.4352
2003	0.2493	0.305293794	0.6388	−0.0560	−0.3895
2004	0.4900	0.280942956	0.7031	0.2091	−0.2131
2005	4.2747	4.273616143	0.722	0.0011	3.5527
2006	0.1208	0.131325997	0.3682	−0.0105	−0.2474
2007	0.1574	−0.046409431	−0.4118	0.2038	0.5692
2008	−0.4903	−0.062363391	−0.2475	−0.4279	−0.2428
2009	−0.2563	−0.011917254	−0.0344	−0.2444	−0.2219
2010	0.2303	0.07103503	0.2881	0.1593	−0.0578
2011	0.2559	0.167068188	0.498	0.0888	−0.2421
2012	0.5182	0.267978439	0.5949	0.2502	−0.0767
2013	0.5520	0.370660128	0.6583	0.1813	−0.1063
2014	0.4821	0.473980684	0.6636	0.0081	−0.1815
2015	0.3392	0.577531441	0.6742	−0.2383	−0.3350
2016	0.7254	0.681165131	0.7208	0.044235	0.0046
2017	1.1472	0.784828629	0.6572	0.362371	0.49
2018	0.5684	0.888502772	0.7105	−0.3201	−0.1421
2019	0.8025	0.992180662	0.668	−0.18968	0.1345

Abbreviations: ANFIS, adaptive network-based fuzzy inference system; ARIMA, autoregressive integrated moving average.

Figure 6 shows that the outputs of ANFIS follow the direction of actual data, whereas the ARIMA has not.

5 | CONCLUSION

The LR is an important ratio related to the efficiency of all functions utilized by insurance companies. As such, in this paper, we showed the ability of ANFIS and ARIMA models to predict the LR for the petroleum insurance branch in Egypt's most significant national insurance company (Misr Insurance Holding Company). The LR data was obtained from the annual report of the Egyptian Financial Supervisory Authority for insurance companies. All data used to analyze the ANFIS and ARIMA models' performance in this study was collected from 25 years of data, ranging from 1995 to 2019. Moreover, we utilized the ANFIS model for one input and one output. We utilized 85% of the data for training the model and the remaining 15% of the data to create testing for the model. The same data was utilized for the ARIMA model.

Our study results show that the ANFIS model outperforms the ARIMA models in predicting the LRs in the petroleum insurance branch. This paper's contribution to existing knowledge includes a mathematical model that can be used to predict an essential ratio in the insurance company. The findings of this study and the implications of the ANFIS model have produced several points that could be areas of interest for future studies, including the prediction of other insurance ratios that may impact the implementation of other policies or its use in describing the insolvency of insurance companies as a mean of providing an early warning signal.

ACKNOWLEDGMENT

The authors would like to thank the respected reviewers for their valuable and helpful comments and suggestions.

ORCID

Ahmed A. Khalil http://orcid.org/0000-0002-4563-5822

REFERENCES

Abbasi, D., Ashrafzadeh, A., Asheghi, R., & Merufinia, E. (2014). A survey on the performance of fuzzy-neural network at predicting the average monthly discharge of catchment basin areas having snow regimes. *Journal of Civil Engineering and Urbanism*, *4*, 480–484.

Abd El-Zahir, A. (2015). Statistical model for predicting loss ratios in property insurance sector in Egyptian insurance market. *Journal of Contemporary Commercial Research, Faculty of Commerce, Sohag University*, *2*, 105–135.

Akkoç, S. (2012). An empirical comparison of conventional techniques, neural networks and the three stage hybrid adaptive neuro fuzzy inference system (ANFIS) model for credit scoring analysis: The case of Turkish credit card data. *European Journal of Operational Research*, *222*(1), 168–178.

Bakhit, S. A., El-Shami, O. M., & Mohamed, H. I. (2004). Management parameters affecting wheat harvesting losses at the Rahad Agricultural Scheme, Sudan. *University of Khartoum Journal of Agricultural Sciences (Sudan)*, *12*(3), 405–414.

Berhe, T. A., & Kaur, J. (2017). Determinants of insurance companies' profitability analysis of insurance sector in Ethiopia. *International Journal of Research in Finance and Marketing (IJRFM)*, *7*(4), 124–137.

Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis: Forecasting and control*. Holden-Day.

Çakıt, E., Karwowski, W., Murata, A., & Olak, A. J. (2020). Application of structural equation modeling (SEM) and an adaptive neuro-fuzzy inference system (ANFIS) for assessment of safety culture: An integrated modeling approach. *Safety*, *6*(1), 14.

Cappiello, A. (2020). Risks and control of insurance undertakings, *The European insurance industry* (pp. 7–29). Springer.

El-Bassiouni, M. Y. (1991). A mixed model for loss ratio analysis. *ASTIN Bulletin: The Journal of the IAA*, *21*(2), 231–238.

Geetha, A., & Nasira, G. M. (2016). Time-series modelling and forecasting: Modelling of rainfall prediction using ARIMA model. *International Journal of Society Systems Science*, *8*(4), 361–372.

Hamid, A., & Mohamed, I. (2015). *Forecasting using fuzzy time series models and comparing with boxes and Jenkins models* (Master's thesis in Statistics), Faculty of Commerce, Mansoura University.

Hogg, R. V., & Klugman, S. A. (2009). *Loss distributions* (p. 249). John Wiley & Sons.

Ibrahim, Y., Tarek, A., & Mohd, S. M. (2011). Cross sectional estimation of loss reserve for Egyptian fire insurance market. *European Journal of Social Sciences*, *25*(4), 456–464.

Jang, J.-S. (1993). ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man and Cybernetics*, *23*(3), 665–685.

Jovanovic, B. B., Reljin, I. S., & Reljin, B. D. (2004, September). *Modified ANFIS architecture-improving efficiency of ANFIS technique*. 7th Seminar on Neural Network Applications in Electrical Engineering, 2004 (NEUREL 2004), IEEE, pp. 215–220.

Kasturi, R. (2006). Performance management in insurance corporation. *Journal of Business Administration Online*, *5*(1), 1–15.

Kruse, R. (2008). Fuzzy neural network. *Scholarpedia*, *3*(11), 6043.

Lamm-Tennant, J., Starks, L. T., & Stokes, L. (1992). An empirical Bayes approach to estimating loss ratios. *Journal of Risk and Insurance*, *59*, 426–442.

Lin, J.-Y., Cheng, C.-T., & Chau, K.-W. (2006). Using support vector machines for long-term discharge prediction. *Hydrological Sciences Journal*, *51*(4), 599–612.

Nauck, D., Klawonn, F., & Kruse, R. (1997). *Foundations of neuro-fuzzy systems*. John Wiley & Sons, Inc.

Nayak, P. C., Sudheer, K. P., Rangan, D. M., & Ramasastri, K. S. (2004). A neuro-fuzzy computing technique for modeling hydrological time series. *Journal of Hydrology*, *291*(1–2), 52–66.

Oroian, M. (2015). Influence of temperature, frequency and moisture content on honey viscoelastic parameters–Neural networks and adaptive neuro-fuzzy inference system prediction. *LWT-Food Science and Technology*, *63*(2), 1309–1316.

Safwat, H. (2003). Using time series for loss ratios analysis of general insurances portfolio for Egyptian insurance companies. *Egyptian Journal for Commercial Studies, Faculty of Commerce, Mansoura University*, *27*(2). 227–259.

Soliman, O. R. (2003). *Forecasting loss ratio in property and liability insurance companies using autoregressive and integrated moving average models (ARIMA) for time series analysis* (Vol. 1). New Afaq, Faculty of Commerce, Menoufia University.

Vieira, J., Dias, F. M., & Mota, A. (2004). Neuro-fuzzy systems: A survey, *5th WSEAS NNA International Conference on Neural Networks and Applications*, March 2004, Udine, Italia, pp. 87–92.

Yunos, Z. M., Shamsuddin, S. M., Sallehuddin, R., & Alwee, R. (2019). Hybrid predictive modelling for motor insurance claim. *IOP Conference Series: Materials Science and Engineering*, *551*(1), 12075.

How to cite this article: Khalil, A. A., Liu, Z., & Ali, A. A. (2022). Using an adaptive network-based fuzzy inference system model to predict the loss ratio of petroleum insurance in Egypt. *Risk Management and Insurance Review*, 1–14. https://doi.org/10.1111/rmir.12200