ISSN
2338-0128

PROCEEDINGS BOOK
THE 8th ANNUAL BASIC SCIENCE INTERNATIONAL CONFERENCE

6-7 March 2018
Ijen Suites Hotel Resort & Convention
Malang, Indonesia

Convergence of Basic Sciences, Toward the World's Sustainability Challenges

Topics
Energy
Environmental Science and Technology
Instrumentation and Measurement
Material Science and Technology
Mathematics, Statistics, and Modelling
Molecular and Health Science
Science and Technology Education

Time Series Analysis of Claims Reserve in General Insurance Industry

Mujiati Dwi Kartikasari[1,a)] and Nurul Imani[1]

[1] *Department of Statistics, Universitas Islam Indonesia*
Jalan Kaliurang Km 14.5 Sleman, Yogyakarta, Indonesia 55584

[a)]Corresponding author: mujiatikartikasari@uii.ac.id

Abstract. This paper presents a different approach for reserving of claims in general insurance industry that is using time series analysis. The authors of this paper propose deterministic time series models such as moving average and exponential smoothing to predict claims reserve. Then the stochastic model, i.e. autoregressive integrated moving average (ARIMA), is appeared as a comparator for the deterministic models. The proposed models are applied to monthly liability claims data of a general insurance company in Indonesia. The comparison shows that the ARIMA of stochastic model is preferred.

INTRODUCTION

To be able to fulfill future obligations, general insurance companies set claims reserve routinely at the moment evaluation. Generally, prediction of claims reserve in actuarial literature is based on scheme of run-off triangle data [1,2,3] as shown in **TABLE 1**.

TABLE 1. Scheme of run-off triangle data.

Claims Occurrence Period	Development Period						
	1	**2**	$\cdots$	$\boldsymbol{j}$	$\cdots$	$\boldsymbol{n-1}$	$\boldsymbol{n}$
1	y_{11}	y_{12}	$\cdots$	y_{1j}	$\cdots$	$y_{1(n-1)}$	y_{1n}
2	y_{21}	y_{22}	$\cdots$	y_{2j}	$\cdots$	$y_{2(n-1)}$	
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$		
$\boldsymbol{i}$	y_{i1}	y_{i2}	$\cdots$	y_{ij}			
$\cdots$	$\cdots$	$\cdots$	$\cdots$				
$\boldsymbol{n-1}$	$y_{(n-1)1}$	$y_{(n-1)2}$					
$\boldsymbol{n}$	y_{n1}						

In scheme of run-off triangle data, historical claims data, y_{ij}, are classified by claims occurrence period $i(i = 1,2,\cdots,n)$ and development period $j(j = 1,2,\cdots,n)$. To obtain prediction of future claims reserve, a calculation of historical claims data using various methods in actuarial literature is used. The prediction of claims reserve will be placed on blank cell in scheme of run-off triangle data.

Although the use of scheme of run-off triangle data is widely applied to predict claims reserve in actuarial literature, there are quite a few problems in the scheme of run-off triangle data. The problems of the scheme of run-off triangle data are discussed in Mutaqin et. al [4]. To that end, this paper suggests a claims reserve prediction using non-actuarial approach specifically time series analysis. We start with deterministic models like moving average and exponential smoothing. Then the stochastic model, autoregressive integrated moving average (ARIMA), is used as a comparator of the deterministic models. By using liability claims data from a general insurance company in Indonesia, we apply the proposed models. Three criteria of measurement error, i.e. root mean squared error (RMSE), mean absolute deviation (MAD), and mean absolute percentage error (MAPE), are used to evaluate the performance of the proposed models.

TIME SERIES MODELS

This section explain the proposed time series models used for predicting claims reserve. The models consist of moving average, exponential smoothing, and ARIMA model.

Moving Average

Moving average is a model with the concept of smoothing historical time series data. One of types of moving average models is double moving average that is model for handling trend. Formulation of the double moving average $MA(M \times N)$ is described as follows [5]:

$$s_t' = \frac{\left(y_t + y_{t-1} + y_{t-2} + \cdots + y_{t-N+1}\right)}{N}, \tag{ix}$$

$$s_t'' = \frac{\left(s_t' + s_{t-1}' + s_{t-2}' + \cdots + s_{t-M+1}'\right)}{M}, \tag{x}$$

$$a_t = s_t' + \left(s_t' - s_t''\right) = 2s_t' - s_t'', \tag{xi}$$

$$b_t = \frac{2}{M-1}\left(s_t' - s_t''\right), \tag{xii}$$

$$\hat{y}_{t+m} = a_t + b_t(m). \tag{xiii}$$

The notation of s_t' and s_t'' denote the simple moving average and the double moving average at the time period t, respectively. The notation of a_t is the single moving average adjustment at the time period t, b_t is the trend estimate, and $\hat{y}_{t+m}$ is prediction value for m future periods.

Exponential Smoothing

Exponential smoothing is a model for smoothing historical time series data using an exponentially weighted smoother. Holt's exponential smoothing is the one of exponential smoothing models based on Holt's method. This method divides the time series data into two components, i.e. the level, a_t, and the trend, b_t, that can be calculated from [6]:

$$a_t = \alpha y_t + (1-\alpha)(a_{t-1} + b_{t-1}), \quad 0 < \alpha < 1, \tag{(1}$$

$$b_t = \beta(a_t - a_{t-1}) + (1-\beta)b_{t-1}, \quad 0 < \beta < 1, \tag{(2}$$

with the initial value $a_1 = y_1$ and $b_1 = y_2 - y_1$. Furthermore, the sum of the level and trend components at the time period t can be used as prediction of m future periods that can be formulated by:

$$\hat{y}_{t+m} = a_t + b_t(m). \tag{(3}$$

ARIMA

ARIMA model is a model which uses a number of observations for several past periods as a basis for the preparation of a prediction for the future. This model can describe the effect of autoregressive (AR) and moving average (MA). The general form of ARIMA(p, d, q) model is [6]:

$$\Phi(B)(1-B)^d y_t = \delta + \Theta(B)\varepsilon_t, \tag{(4}$$

with $\Phi(B) = \left(1 - \varphi_1 B - \varphi_2 B^2 - \cdots - \varphi_p B^p\right)$ and $\Theta(B) = \left(1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q\right)$. The notation B is backshift operator and ε_t is a white noise process with zero mean and constant variance.

RMSE, MAD, and MAPE

Performance of time series models can be evaluated by using many kinds of measurement error. In this paper, we use three criteria of measurement error, that are RMSE, MAD, and MAPE. The formula of the RMSE, the MAD, and the MAPE are described as follows [6]:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(y_t - \hat{y}_t\right)^2}\ , \tag{(5)}$$

$$MAD = \frac{1}{n}\sum_{t=1}^{n}\left|y_t - \hat{y}_t\right|, \tag{(6)}$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right| \times 100\%\,. \tag{(7)}$$

RESULTS

The above time series models are applied to monthly liability claims data from an Indonesian general insurance company, with observation period January 2009 until December 2014. The time series data of the liability claims show growth trend although there is the sharp drop in the end of 2012 as shown in **FIGURE 1**. The data from January 2009 up to and including December 2013 are used as in-sample data to predict the 2014 data as out-of-sample data. The differences between the actual and the prediction for the 12 months of 2014 are used to evaluate the proposed models performance.

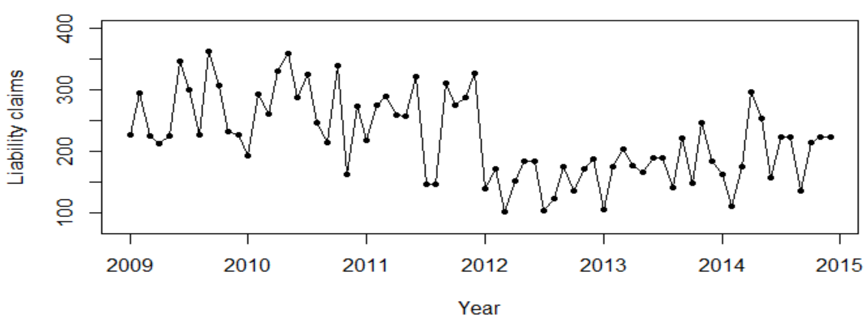

FIGURE 1. Plot of liability claims data.

The double moving average, Holt's exponential smoothing, and ARIMA are implemented on the liability claims data to predict claims reserve in the future. Analysis of the data by using the double moving average and the Holt's exponential smoothing tend to be easier to perform than the ARIMA. The ARIMA model have to saturate the assumption of independent residual and all parameters should be significant. By using Box-Jenkins procedure, the ARIMA model for the data is:

$$\left(1 - 0.3138B - 0.2724B^2\right)y_t = 226.4886 + \varepsilon_t\,. \tag{xiv}$$

FIGURE 2 shows the prediction results as obtained from the double moving average, Holt's exponential smoothing, and ARIMA model. The plots show that the prediction of Holt's exponential smoothing and ARIMA model are close from the actual data both in the in-sample data and out-of-sample data.

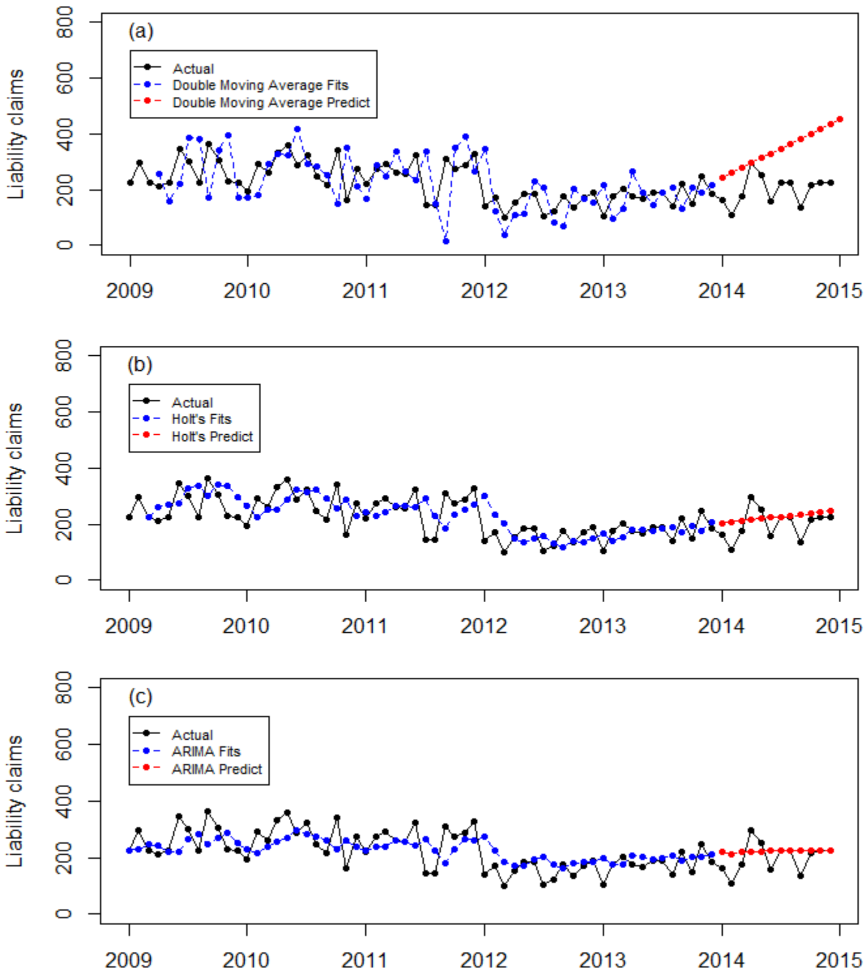

FIGURE 2. Claims reserve prediction by (a) double moving average, (b) Holt's exponential smoothing, (c) ARIMA.

To demonstrate the performance of the double moving average, Holt's exponential smoothing, and ARIMA accurately both in the in-sample and out-of-sample data, the authors use the three measurement error criteria include RMSE, MAD, and MAPE. The results are shown in **TABLE 2**.

TABLE 2. RMSE, MAD, and MAPE comparison.

Data	Measurement Error	Model		
		Double Moving Average	Holt's Exponential Smoothing	ARIMA
In-sample	RMSE	97.29918	313.4959	62.49371
	MAD	77.07403	74.63681	48.37917
	MAPE	37.77467	33.23954	24.55954
Out-of-sample	RMSE	164.7118	56.48467	55.33709
	MAD	148.8837	46.50797	43.12226
	MAPE	83.67315	28.58931	27.22728

Measurement error of the in-sample data in **TABLE 2** indicate that the ARIMA model outperformed the double moving average and Holt's exponential smoothing model for predicting claims reserve. Likewise, for out-of-sample prediction of the data, ARIMA model outperforms the other models. Therefore, to predict claims reserve in general insurance industry, ARIMA model can be a good alternative.

SUMMARY

We have compared deterministic time series models, i.e. double moving average and Holt's exponential smoothing, and stochastic time series model, i.e. ARIMA. Based on the measurement error comparison, ARIMA model outperformed the other models. Based on the results, ARIMA of stochastic model could be a good alternative to predict claims reserve in general insurance industry because it performed well both in the in-sample and out-of-sample data.

ACKNOWLEDGMENTS

This research is funded by Directorate of Research and Community Service Universitas Islam Indonesia. Thanks to Hafizah Ilma for her help in presenting this result in the 8th Annual Basic Science International Conference (BaSIC), Malang-Indonesia.

REFERENCES

1. T. Mack, ASTIN Bulletin **23**, 213 – 225 (1993).
2. P. J. R. Pinheiro, J. M. A. E Silva, and M. D. L. Centeno, Journal of Risk and Insurance **70**, 701 – 714 (2003).
3. K. Antonio, J. Beirlant, T. Hoedemakers, and R. Verlaak, North American Actuarial Journal **10**, 30 – 48 (2006).

A. K. Mutaqin, D. R. Tampubolon, and S. Darwis, Statistika **8**, 55 – 59 (2008).

4. S. Makridakis, S. C. Wheelwright, and V. E. McGee, *Forecasting: Methods and Applications Second Edition* (John Wiley & Sons Inc, Canada, 1983).
5. D. C. Montgomery, C. L. Jennings, and M. Kulahci, *Introduction to Time Series Analysis and Forecasting Second Edition* (John Wiley & Sons Inc, New Jersey, 2015).