

Research Article

Prediction of Outstanding Claims Liability in Non-Life Insurance: An Application of Adaptive Grey Model

Mujiati Dwi Kartikasai^{1*}, Hani'atul Maghfuroh¹

¹Statistics Department FMIPA Universitas Islam Indonesia, Jln. Kaliurang Km 14.5 Sleman, Yogyakarta, Indonesia

*Corresponding author: mujiatikartikasari@uii.ac.id

Received: 25 September 2020; Accepted: 4 September 2021; Published: 16 September 2021

Abstract: In order to assess the solvency of non-life insurance companies, the prediction of outstanding claims liability is very important. Prediction of outstanding claims liability is usually done by using a run-off triangle data scheme. However, if data are not available to form the scheme, the prediction of outstanding claims liability cannot be made. Another alternative for predicting of outstanding claims liability is to use time series analysis. This research uses an adaptive grey model that has the advantage of being free of assumptions of data patterns and a minimum amount of data used to predict is small (at least 4 data). To determine the accuracy of the adaptive grey model, we compare the prediction of outstanding claims liability using a grey model classic. Based on the analysis results, the adaptive grey model is better than the classic grey model in predicting outstanding claims liability.

Keywords: Outstanding Claims Liability, Adaptive Grey Model, Non-Life Insurance, Time Series

Introduction

In non-life insurance, payment of claims that have occurred and reported to the company sometimes requires quite a long time to settle. This causes companies to have obligations of claims which are not finalized that must be fulfilled immediately. In insurance, the obligations are often referred to as outstanding claims liability (OCL). OCL prediction is very important to do, because this information is very important for insurance companies in assessing solvency (the ability of an insurance company to fulfill its obligations).

Generally, prediction of OCL is performed using run-off triangle data scheme [1-5]. The run-off triangle scheme is a data scheme that contains an aggregate claim description and is a summary of the individual claim data set [6]. In this scheme, information from the observed claim data is formed into an upper triangular matrix of size $N \times N$ with the occurrence period as a column and the development period of payment as a row. However, if information about the observed claim data is not available then a run-off triangle data scheme cannot be formed. Consequently, OCL predictions cannot be made.

Another alternative that can be used to predict OCL is to use time series analysis. Time series analysis is an analysis that is used to make predictions by utilizing data collected in a series of particular time periods or intervals. However, the use of classical time series analysis both deterministic and stochastic often has several assumptions that must be fulfilled, such as time series data patterns and minimum amount of data used. For this reason, this paper uses a grey model that does not require the assumption of data patterns and prediction can be made even though the data is very limited (the least 4 data) [7].

The grey model is part of the grey system theory which was introduced by Deng Julong in 1982 [8]. Because of its ease of use, the grey model is widely applied in various fields. For example, in the telecommunications [7], healthcare [9-14], electricity demand [15-18], energy problems [19-20], transportation [21-22], economic fields [23-24], and so on. The basic form of the grey model is GM(1,1). Although GM(1,1) has been widely applied with good prediction accuracy results, this model still has limitations in its use, because it cannot reflect of data trends at different period of time. To solve this problem, the trend and potential tracking method (TPTM) proposed by Li and Yeh (2008) [25] is used in

GM(1,1) to construct an adaptive grey model which is AGM(1,1). TPTM is useful for analyzing data behavior, extracting hidden information from the data, and utilizing the trend and potency value to build the AGM(1,1) [25] so that the accuracy of prediction is better than GM(1,1) [26].

Materials and Methods

Materials

Data used in this paper are annual data of OCL obtained from Capital Market and Financial Institutions Supervisory Agency [27-31] and Indonesia Financial Services Authority [32-37]. The data consist of fourteen observations and are presented in Table 1.

Table 1. Data of OCL

Year	OCL (Million Rp)	Year	OCL (Million Rp)
2002	642,793	2011	946,807
2003	505,311	2012	1,169,461
2004	378,523	2013	1,212,360
2005	455,009	2014	1,957,155
2006	520,920	2015	1,625,891
2007	555,125	2016	2,017,003
2008	591,752	2017	2,037,826
2009	782,188	2018	2,468,932
2010	704,436	2019	2,495,228

Trend and Potency Tracking Methods

TPTM is an analytical method that uses data characteristics to explore possible changes in data behavior at different period of time [26]. TPTM is the key used in AGM(1,1) to improve accuracy of GM(1,1).

The procedure in determining TPTM which is cited from Chang et.al (2013) are as follows [26].

Step 1: Arrange series data $X = \{x_1, x_2, \dots, x_n\}$. Let x_{\min} be the minimum value of X and x_{\max} be the maximum value of X .

Step 2: Determine the variations σ_i of paired data (x_{i-1}, x_i) for $i = 2, 3, \dots, n$. Next, determine increasing or decreasing potencies according to sequence of the data. If σ_i is positive, then the trend of data in the time period i increases, and vice versa.

Step 3: Define weights to the latest data. The weight of data in the time period i , denoted by w_i , equals $w_i = i - 1$ for $i = 2, 3, \dots, n$.

Step 4: Compute $A_i = \sigma_i \times w_i$ for $i = 2, 3, \dots, n$. The use of A_i is as operator to strengthen the data trend and potency for each time period. If $A_i > 0$, then there is increasing potency (IP), and if $A_i < 0$, then there is decreasing potency (DP).

Step 5: Compute the central location (CL) using the equation $CL = \frac{x_{\min} + x_{\max}}{2}$. CL is used as the main point to expand the data domain range.

Step 6: Calculate the average of increasing potency (AIP) and average of decreasing potential (ADP), then use both to expand the range of data domains. The upper limit of the extended domain range (EDR_UL) is obtained by the equation $EDR_UL = x_{\max} + AIP$ and the lower limit of the extended domain range (EDR_LL) is obtained by the equation $EDR_LL = x_{\min} + ADP$.

Step 7: Form a triangular TP function using the CL, EDR_UL, and EDR_LL. Compute TP value from the corresponding data by using triangle rule. For CL, set the TP value of CL to 1. The following is an example of a triangle for calculating the TP value presented in Figure 1.

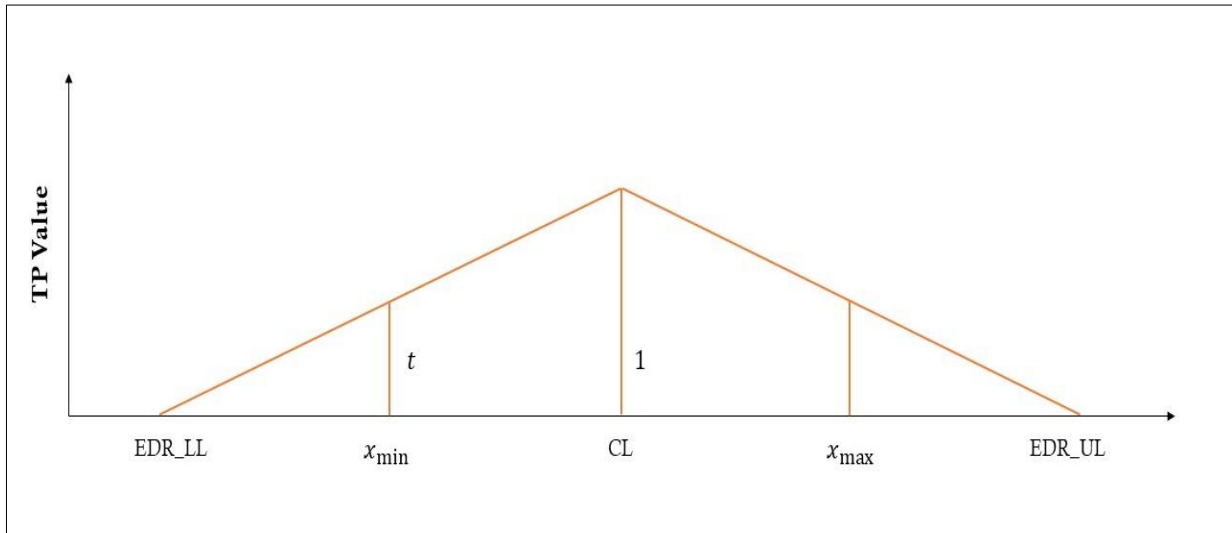


Figure 1. An example of a triangle for calculating the TP value [9]

Adaptive Grey Model

AGM(1,1) is a model obtained by integrating TPTM value into the modeling process of GM(1,1). The procedure of AGM(1,1) is described as follows [26].

Step 1: Arrange an original data series $X^{(0)}$ as follows:

$$X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\}, n \geq 4. \quad (1)$$

Step 2: Compute TP values by using the TPTM and arrange TP values as $\{TP_i\} = \{TP_1, TP_2, \dots, TP_n\}, i = 1, 2, \dots, n$.

Step 3: Calculate the coefficient α_k with the following equation

$$\alpha_k = \frac{\sum_{i=1}^k z^{i-1} TP_i}{\sum_{i=1}^k z^{i-1}}, k \geq 2. \quad (2)$$

Step 4: Create a new data series which is the accumulation of k original data series through accumulating generation operator (AGO), which is,

$$X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)\} \quad (3)$$

where $x^{(1)}(1) = x^{(0)}(1)$ and $x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i), k = 2, 3, \dots, n$.

Step 5: Calculate the background values $z^{(1)}(k)$. The $z^{(1)}(k)$ is expressed as

$$z^{(1)}(k) = (1 - \alpha_k)x^{(1)}(k - 1) + \alpha_k x^{(0)}(k); \quad (4)$$

where $\alpha \in (0,1)$ and $k = 2, 3, \dots, n$.

Step 6: Construct differential equation from (1) and (4), that is,

$$x^{(0)}(k) + az^{(1)}(k) = b, \quad (5)$$

where a and b are the developing coefficient and the grey input coefficient, respectively. By using the least square method, the estimation of a and b can be obtained, that is

$$\begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix} = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix} \times \begin{bmatrix} a \\ b \end{bmatrix}. \quad (6)$$

Let

$$Y = [x^{(0)}(2), x^{(0)}(3), \dots, x^{(0)}(n)]^T,$$

$$\hat{a} = [a, b]^T,$$

$$B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}, \quad (7)$$

$$\hat{a} = (B^T B)^{-1} B^T Y.$$

Step 7: Calculate the prediction value in period $k + 1$ with an initial value of $x^{(0)}(1) = x^{(1)}(1)$ with the equation

$$\begin{aligned} x^{(1)}(k + 1) &= \left(x^{(0)}(1) - \frac{b}{a}\right) e^{-ak} + \frac{b}{a}, \\ \hat{x}^{(0)}(k + 1) &= \hat{x}^{(1)}(k + 1) - \hat{x}^{(1)}(k). \end{aligned} \quad (8)$$

Model Evaluation

Good prediction accuracy is indicated by the minimal prediction error value. Prediction error can be used for two situation. The first is making a choice between several alternative prediction models, and the other is evaluating the success or failure of the models used. In this paper, we will use mean absolute error (MAE) and mean absolute percentage error (MAPE) as a prediction error measure. The MAE and MAPE are defined as follows [38].

$$MAE = \frac{1}{n} \sum_{i=1}^n |x^{(0)}(i) - \hat{x}^{(0)}(i)|. \quad (9)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{x^{(0)}(i) - \hat{x}^{(0)}(i)}{x^{(0)}(i)} \right| \times 100\%. \quad (10)$$

Results and Discussions

The OCL data presented in Table 1 is divided into two parts, namely training data and testing data. Training data is used to form models and testing data is used to evaluate models. Data from 2002 to 2005 is used as training data, and data from 2006 to 2019 is used as testing data.

By using AGM (1,1), the training data from 2002 to 2005 will be used to build the model. The calculation process is shown in the following steps, and the results are summarized in Table 2. Meanwhile, the summary of the OCL prediction results is presented in Table 3.

1. Arrange an original data series $X = \{642,793; 505,311; 378,523; 455,009\}$.
2. Compute TP values using TPTM. We get $TP = \{0.6346, 0.9837, 0.5967, 0.8302\}$.
3. Calculate the coefficient $\alpha_k = \{0.8673, 0.7127, 0.7753\}$.
4. Create a new data series $X^{(1)}$ using AGO. We get $X^{(1)} = \{642,793; 1,148,104; 1,526,627; 1,981,636\}$.
5. Calculate the background values $Z^{(1)} = \{1,081,056; 1,417,876; 1,879,417\}$.
6. Calculate $\hat{a} = (B^T B)^{-1} B^T Y$, where $Y = [505,311; 378,523; 455,009]^T$ and $B = \begin{bmatrix} -1,081,056 & 1 \\ -1,417,876 & 1 \\ -1,879,417 & 1 \end{bmatrix}$.
We get $\hat{a} = [0.0493; 518,299.9]^T$.
7. Use (8) to build the model $x^{(1)}(k + 1) = -9,860,461.53 e^{-0.0493k} + 10,503,254.53$. We get the prediction value $\hat{x}_5 = 409,439.93$.

Table 2. Calculation Results of AGM(1,1)

Year	x_k	TP Value	$x_k^{(1)}$	α_k	$Z_k^{(1)}$	\hat{x}_k
2002	642,793	0.6346	642,793	-	-	-
2003	505,311	0.9837	1,148,104	0.8673	1,081,056	474,769.78
2004	378,523	0.5967	1,526,627	0.7127	1,417,876	451,910.16
2005	455,009	0.8302	1,981,636	0.7753	1,879,417	430,151.21
2006	520,920	-	-	-	-	409,439.93

Table 3. Prediction Output of AGM(1,1)

Year	Actual Values	Prediction Values
2002	642,793	
2003	505,311	
2004	378,523	
2005	455,009	
2006	520,920	409,439.93
2007	555,125	569,048.06
2008	591,752	608,598.19
2009	782,188	622,867.63
2010	704,436	896,037.18
2011	946,807	781,719.47
2012	1,169,461	967,073.28
2013	1,212,360	1,360,201.44
2014	1,957,155	1,381,082.10
2015	1,625,891	2,441,139.90
2016	2,017,003	1,910,407.19
2017	2,037,826	1,929,100.29
2018	2,468,932	2,168,157.86
2019	2,495,228	2,707,632.72

To test the effectiveness of the AGM (1,1) model, we compare the results of prediction value with GM(1,1) model. The prediction results are presented in Figure 2.

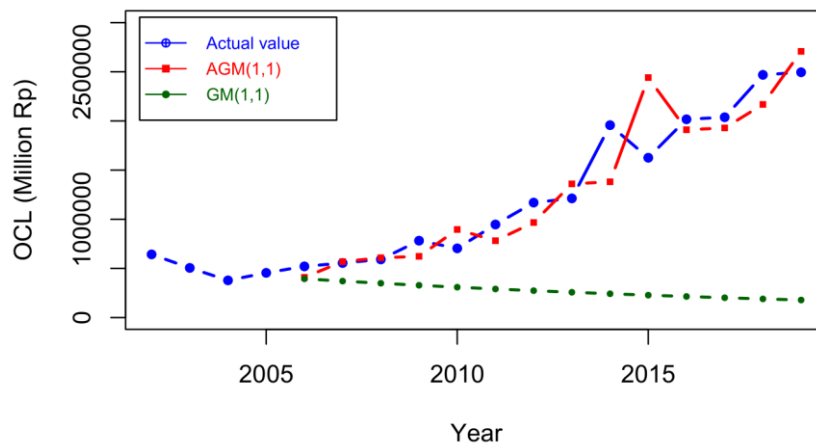


Figure 2. Prediction results using AGM(1,1) and GM(1,1)

Figure 2 shows that the prediction results using AGM(1,1) is closer to the actual data and have the same pattern as the actual data. Meanwhile, the prediction results using GM(1,1) show the opposite. To see a good prediction accuracy, a measure of model evaluation based on MAE and MAPE is calculated. The results of comparing OCL predictions based on MAE and MAPE using AGM(1,1) and GM(1,1) are presented in Table 4.

Table 4. Comparison of AGM(1,1) and GM(1,1)

Model	MAE	MAPE (%)
AGM(1,1)	223,450.70	16.58
GM(1,1)	1,089,594.41	69.66

Based on the results shown in Table 4, MAE and MAPE of the AGM(1.1) are lower than MAE and MAPE of the GM(1.1). So, it can be said that the AGM(1,1) gives better results than the GM(1,1) in predicting OCL. Furthermore, the AGM(1,1) can be an alternative model in predicting OCL.

Conclusion

Grey model can be utilized as an alternative method that can be used to predict OCL besides using the run-off triangle data scheme. In this study, the AGM(1,1) model was used to predict OCL. To verify the effectiveness of the AGM(1,1), the prediction value of the GM(1,1) model was used as a comparison. Based on the results and discussions, the AGM(1,1) has smaller value of MAE and MAPE compared to GM(1,1) in predicting OCL in the non-life insurance industry.

References

- [1] R. Kaas, J. Dhaene, M. Goovaerts, M. Denuit, *Modern Actuarial Risk Theory using R* Second Edition, Springer 2008.
- [2] M.V. Wuthrich, M. Merz, *Stochastic Claims Reserving Methods in Insurance*, John Wiley & Sons 2008.
- [3] I. Badounas, G. Pitselis, *Loss Reserving Estimation With Correlated Run-Off Triangles in a Quantile Longitudinal Model*, *Risks* 8(14) (2020).
- [4] F. F. Adam, *Claim Reserving Estimation by Using the Chain Ladder Method*, *KnE Social Sciences* (2018) 1192-1204.
- [5] R. N. Suwandani, Y. Purwono, *Implementation of Gaussian Process Regression in Estimating Reserve for Motor Vehicle Insurance Claim*, *Journal of Asian Multicultural Research for Economy and Management Study* 2(1) (2021) 38-48.
- [6] K. Antonio, J. Beirlant, T. Hoedemakers, R. Verlaak, *Lognormal Mixed Models for Reported Claims Reserves*, *North American Actuarial Journal* 10(1) (2006) 30-48.
- [7] C. Lin, *Forecasting and Analyzing the Competitive Diffusion of Mobile Cellular Broadband and Fixed Broadband in Taiwan with Limited Historical Data*, *Economic Modelling* 35 (2013) 207-213.
- [8] D. Julong, *Introduction to Grey System Theory*, *The Journal of Grey System* 1 (1989) 1-24.
- [9] H. Dang, Y. Huang, C. Wang, T. Nguyen, *An Application of the Short-Term Forecasting with Limited Data in the Healthcare Traveling Industry*, *Sustainability* 8 (2016) 1037-1051.
- [10] C. Zor, F. Cebi, *Demand prediction in health sector using fuzzy grey forecasting*, *Journal of Enterprise Information Management* 31(1) (2018).
- [11] A. D. L. Foucade, S. Gabriel, E. Scott, K. Theodore, C. Metivier, *A Survey of Selected Grey Forecasting Models with Application to Medical Tourism Forecasting*, *Theoretical Economics Letters* 9(4) (2019) 1079-1092.
- [12] Y. Zhao, M. Shou, Z. W. Prediction of the Number Patients Infected with COVID-19 Based on Rolling Grey Verhulst Models, *International Journal of Environmental Research and Public Health* 17(12) (2020).
- [13] X. Yang, J. Zou, D. Kong, G. Jiang, *The analysis of GM (1, 1) grey model to predict the incidence trend of typhoid and paratyphoid fevers in Wuhan City, China*, *Medicine* 97(34) (2018).
- [14] B. M. Iqelan, *Forecasts Breast Cancer Referrals Using Grey Prediction Model GM(1,1)*, *Applied Mathematical Sciences* 11(54) (2017) 2647-2662.
- [15] K. Li, T. Zhang, *Forecasting Electricity Consumption Using an Improved Grey Prediction Model*, *Information* 9(8) (2018) 204.
- [16] S. Guefano, J. G. Tamba, T. E. W. Azong, L. Monkam, *Methodology for Forecasting Electricity Consumption by Grey and Vector Autoregressive Models*, *MethodsX* 8 (2021).
- [17] F. Song, J. Liu, T. Zhang, J. Guo, S. Tian, D. Xiong, *The Grey Forecasting Model for the Medium- and Long-Term Load Forecasting*, *Journal of Physics: Conference Series* 1654 (2020).
- [18] L. Zhao, X. Zhou, *Forecasting Electricity Demand Using a New Grey Prediction Model with Smoothness Operator*, *Symmetry* 10 (2018).
- [19] B. Zeng, C. Li, *Forecasting the Natural Gas Demand in China Using a Self-Adapting Intelligent Grey Model*, *Energy* 112 (2016) 810-825.
- [20] N. geng, J. Li, H. Gao, S. Zhao, *Forecasting Hydrogen Energy Production in China base on Multi Grey Model*, *IOP Conference Series: Earth and Environmental Science* 811 (2021).
- [21] F. Pourofoghi, J. S. Ardabili, D. D. Salokolaei, *A New Approach for Finding an Optimal Solution for Grey Transportation Problem*, *International Journal of Nonlinear Analysis and Applications* 10 (2019).
- [22] C. Liu, T. Shu, S. Chen, S. Wang, K. K. Lai, L. Gan, *An improved grey neural network model for predicting transportation disruptions*, *Expert Systems with Application* 45 (2016) 331-340.
- [23] H. Cy, L. Cy, C. Ci, *Application of Grey Theory to the Field of Economic Forecasting*, *Journal of Global Economics* 3(1) (2014).

-
- [24] E. Kayacan, B. Ulutas, O. Kaynak, Grey System Theory-Based Models in Time Series Prediction, *Expert Systems with Applications* 37 (2010) 1784-1789.
- [25] D. Li, C. Yeh, C. Chang, An Improved Grey-Based Approach for Early Manufacturing Data Forecasting, *Computers & Industrial Engineering* 57 (2009) 1161-1167.
- [26] C. Chang, D. Li, W. Dai, C. Chen, Utilizing an Adaptive Grey Model for Short-Term Time Series Forecasting: A Case Study of Wafer-Level Packaging, *Hindawi Publishing Corporation Mathematical Problems in Engineering* 2013 (2013).
- [27] Insurance Bureau, Indonesian Insurance, Capital Market and Financial Institutions Supervisory Agency 2006.
- [28] Insurance Bureau, Indonesian Insurance, Capital Market and Financial Institutions Supervisory Agency 2007.
- [29] Insurance Bureau, Indonesian Insurance, Capital Market and Financial Institutions Supervisory Agency 2008.
- [30] Insurance Bureau, Indonesian Insurance, Capital Market and Financial Institutions Supervisory Agency 2009.
- [31] Insurance Bureau, Indonesian Insurance, Capital Market and Financial Institutions Supervisory Agency 2010.
- [32] OJK, Insurance Statistics, Financial Services Authority 2014.
- [33] OJK, Insurance Statistics, Financial Services Authority 2015.
- [34] OJK, Insurance Statistics, Financial Services Authority 2016.
- [35] OJK, Insurance Statistics, Financial Services Authority 2017.
- [36] OJK, Insurance Statistics, Financial Services Authority 2018.
- [37] OJK, Insurance Statistics, Financial Services Authority 2019.
- [38] D.C. Montgomery, C.L. Jennings, M. Kulahci, *Introduction to Time Series Analysis and Forecasting*, John Wiley & Sons 2015.