ENTHUSIASTIC
INTERNATIONAL JOURNAL OF APPLIED STATISTICS AND DATA SCIENCE
Volume 3, Issue 1, April 2023, pp. 16-24

Forecasting Population Mortality Rates Using Generalized Lee-Carter Model

I Gusti Nyoman Yudi Hartawan [a,1,*], I Gusti Ngurah Pujawan [a,2], Kadek Mardika Pranata [a,3], Kadek Jayanta [a,4]

[a] Faculty of Mathematics and Natural Sciences Universitas Pendidikan Ganesha, Jln. Udayana, Singaraja-Bali, Singaraja and 8111, Indonesia
[1] yudi.hartawan@undiksha.ac.id*; [2] ngurah.pujawan@undiksha.ac.id; [3] mardikapranata50@gmail.com; [4] jayanta@undiksha.ac.id
* Corresponding author

ARTICLE INFO

Keywords
Mortality rate
Generalized Lee-Carter
ARIMA

ABSTRACT

The demographic process cannot be inseparable from the mortality rate. The appropriate models for forecasting mortality rates are essential in assisting governments, companies, and other agencies in formulating policies or making decisions. As one of the countries with the highest death rate, Japan is influenced by several factors. This research uses the Generalized Lee-Carter Model), which is one of the developments of the Lee-Carter (LC) model. The Lee-Carter model was prevalent by Lee and Carter (1995) as an alternative that is suspected to predict the mortality rate of an area. The first step in this research is to formulate the Generalized Lee-Carter function. Through the function formula, the estimator value of the Generalized Lee-Carter model will be searched in the second stage. And the third stage, through the Generalized Lee-Carter model, will find the RMSE value and then use it in the fourth stage, namely forecasting the future period using ARIMA. The data in this study is facilitated through www.mortality.org, which is one of the Japanese population data. The result of the study showed that the RMSE value for females was 0.01670 and 0.016292 for males. So, it concluded that the Generalized Lee-Carter Model is great for forecasting mortality rates.

1. Introduction

Mortality is one of the three components in the demographic process that influences population structure. Mortality occurs due to age, health, and unpredictable natural factors [1]. The mortality rate is one of the critical factors in influencing the analysis of companies and governments. The mortality rate is presented in the form of a mortality table; the mortality table provides an overview of the life of population groups from birth to death according to time [2]. In forecasting, the mortality table is essential to be paid attention to.

Japan is one of the developed countries that have demography problems. The birth rate in Japan has shown a declining rate, and this is due to the rapid growth of the elderly population. So that this incident resulted in the demographic structure in Japan, showing the condition of an aging society. The low birth has put Japan in the spotlight of some demographic observers [3].

The death rate in Japan is caused by several factors, such as suffering due to illness, economic hardship, family problems, depression, and others [4]. In addition to these factors, many people in Japan choose not to marry. It is due to the increasing needs of life and individual culture [5]. Several Japanese government policies have been implemented to increase the birth rate, but this is not optimal because the death rate in Japan is still high. After all, Japan is ranked third as the country with the cases the highest number of suicides.

The study of mortality rates is an interesting topic to pay attention to. Human life expectancy has increased and decreased in every country. Stochastic is a suitable model for estimating the estimated value and life expectancy [6]. Lee-Carter is one of the good models used in forecasting [7], [8]. Lee and Carter introduced the Lee-Carter model in 1992. In the Lee-Carter model, the mortality index is a matter of concern because it describes the trend from time to time. Still, the Lee-Carter model is limited to past trend patterns, so it does not pay attention to other aspects, such as the environment and health [9]. Based on this, the Lee-Carter model has developed.

The development of the LC model was used to add a factor to reduce the mortality rate [10]. The thing that underlies the difference is the estimated time used in the two models—the factor known. In the GLM approach, it is better able to capture the trend of death by age in the data [11].

In addition to the approach with the GLM model, the cohort effect is added to its development. The addition of this cohort effect provides a better fit and the addition of this cohort effect because the age of the period in the Lee-Carter model does not match the empirical data [12].effect cohort, and the GLM approach used by the Lee-Carter Model becomes generalized to the Generalized Lee-Carter Model. Hence, from this description, researchers will investigate whether the Generalized Lee-Carter Model, which is the general form of the Lee-Carter model, is suitable for forecasting population mortality rates in Japan.

2. Method

This study examines the Generalized Lee Carter model in predicting the death rate of the population. In this case, the expected population death rate in Japan. The research method used in this research is the literature review. In the study, the data used is the death data of the Japanese population from 1990 to 2019. Secondary data is sourced from the *Human Mortality Database*, which can be accessed at www.mortality.org. The data contains the death rate by age, year, and gender. The subjects in this study were residents of Japan from 1990 to 2019. The stages analysis carried out were: 1) Finding an estimator of the Generalized Lee-Carter Model parameters. 2) Finding the error value of the Generalized Lee-Carter model through RMSE by comparing the actual with the estimated results. 3) Forecasting for several future periods with ARIMA.

2.1. Model Generalized Lee-Carter

Renshaw and Huberman first presented this model in a study 2006 as follows:

$$\ln m_{x,t} = \alpha_x + \beta_x^{(1)} k_t + \beta_x^{(0)} \iota_{t-x} \quad (1)$$

where $m_{x,t}$ is the estimated mortality rate at age x and years t, α_x is the average death rate by age over time, $\beta_x^{(0)}$ and $\beta_x^{(1)}$ are parameters that measure the interaction in age with changes in k_t and ι_{t-x}, k_t is the period effect seen in years t, and ι_{t-x} is the cohort effect.

To get the results of each parameter in (1) is done using the approach Poisson and the M model [8], [12], [13]. By using the maximum likelihood estimation of the model parameters M, the updated parameters are obtained as follows.

$$\hat{\alpha}_x^{\{h+1\}} = \alpha_x^{\{h\}} + \frac{\sum_t \omega_{xt}(d_{xt} - \hat{d}_{xt})}{\sum_t \omega_{xt}\hat{d}_{xt}} \quad (2)$$

p-ISSN 2798-253X
e-ISSN 2798-3153

$$\hat{\beta}_x^{(1)\{h+1\}} = \hat{\beta}_x^{(1)\{h\}} + \frac{\sum_t \omega_{xt}\hat{k}_t(d_{xt}-\hat{d}_{xt})}{\sum_t \omega_{xt}\hat{k}_t^2\hat{d}_{xt}} \quad (3)$$

$$\hat{k}_t^{\{h+1\}} = k_t^{\{h\}} + \frac{\sum_x \omega_{xt}\beta_x^{(1)}(d_{xt}-\hat{d}_{xt})}{\sum_x \omega_{xt}\left(\beta_x^{(1)}\right)^2\hat{d}_{xt}} \quad (4)$$

$$\beta_x^{(0)\{h+1\}} = \beta_x^{(0)\{h\}} + \frac{\sum_t \omega_{xt}\hat{\iota}_{t-x}(d_{xt}-\hat{d}_{xt})}{\sum_t \omega_{xt}\hat{\iota}_{t-x}^2\hat{d}_{xt}} \quad (5)$$

$$\hat{\iota}_{t-x}^{\{h+1\}} = \iota_{tx}^{\{h\}} + \frac{\sum_x \omega_{xt}\beta_x^{(0)}(d_{xt}-\hat{d}_{xt})}{\sum_{x,t} \omega_{xt}\left(\beta_x^{(0)}\right)^2\hat{d}_{xt}} \quad (6)$$

By considering several sub-structures of equation (1) by setting $\beta_x^{(0)} = 1$ [14], [15] then (1) changes to (7).

$$ln\, m_{xt} = \alpha_x + \beta_x^{(1)}k_t + \iota_{t-x} \quad (7)$$

So that the model becomes a simpler structure. Forecasting mortality in this model is done using time series estimates, based on the results of the parameters k_t and ι_{t-x} generated using the ARIMA process with the assumption of independence between periods and cohort effects.

3. Results and Discussion

3.1. Results

This study used data from Japan. Data obtained from the site www.mortality.org. the following data were obtained.

Table 1. Japanese Population Death Data

Age	Years						
	1990		1991		…	2019	
	Female	*Male*	*Female*	*Male*	…	*Female*	*Male*
0	2494.51	3128.46	2504.27	2919.39	…	762.39	893.7
1	439.27	548.96	396.2	480.73	…	140.07	138.26
2	221.14	330.57	222.11	292.44	…	77.04	80.15
3	166.1	295.52	156.08	236.36	…	60.03	68.13
4	132.08	236.41	129.06	198.29	…	51.03	51.1
5	133.08	210.37	138.07	167.26	…	46.02	44.09
…	…	…	…	…	…	…	…
96	2310.39	939.63	2489.26	978.47	…	21502	6386.13
97	1613.97	605.05	1674.84	706.06	…	18205.3	4648.83
98	1091.66	397.69	1212.62	431.65	…	14954.6	3429.51
99	657.4	224.39	734.37	264.4	…	11474.9	2458.67
100	460.27	159.28	465.24	164.24	…	7827.99	1497.85

From these data, it can be seen that the death rate of the Japanese population is as shown below. From year to year the death rate of the Japanese population has increased and decreased. From Fig. 1, it can be seen that the male mortality rate is higher than the female from year to year.

p-ISSN 2798-253X
e-ISSN 2798-3153

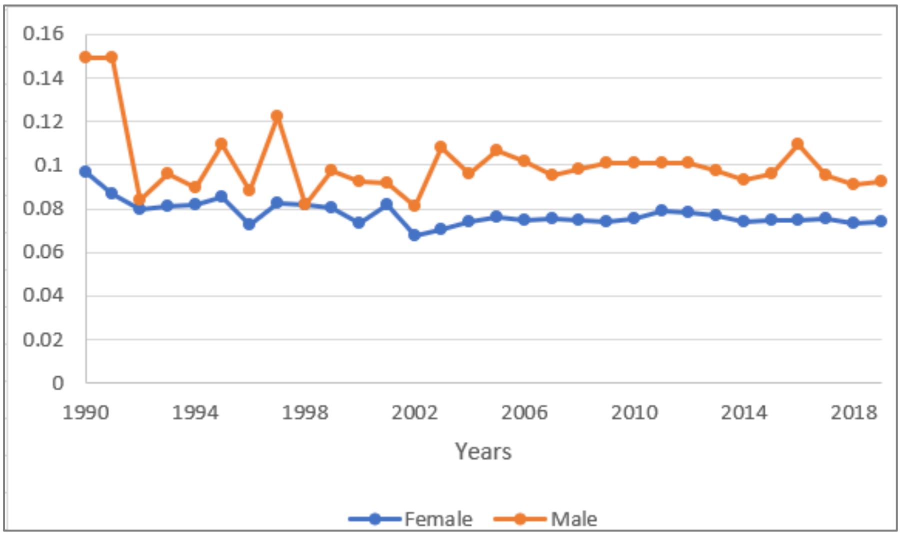

Fig. 1 Death rate by sex.

The parameter estimation results from the Generalized Lee-Carter Model are presented in Table 2 and Table 3:

Table 2. Parameter Estimation of Generalized Lee Carter

Age	Female			Male		
	α_x	$\beta_x^{(1)}$	$\beta_x^{(0)}$	α_x	$\beta_x^{(1)}$	$\beta_x^{(0)}$
0	-5.458177623	0.003522304	1	-5.383879681	0.004283465	1
1	-7.326512785	0.004272541	1	-7.250931184	0.005207919	1
2	-7.911710852	0.002624809	1	-7.835864963	0.003253536	1
3	-8.263254251	0.002410035	1	-8.187213185	0.002998998	1
4	-8.509810681	0.002866741	1	-8.433574761	0.003534735	1
...	...	...	...	...	...	...
97	-0.871855224	0.031470374	1	-1.181137699	0.029488999	1
98	-0.725177558	0.031525295	1	-1.058340799	0.029338795	1
99	-0.581734885	0.031739668	1	-0.939417091	0.029366037	1
100	-0.41250818	0.032536128	1	-0.799725074	0.029992484	1

Table 3. Parameter Estimation of Generalized Lee Carter

Year	Female		Male	
	k_t	ι_{t-x}	k_t	ι_{t-x}
1990	35.95941651	-0.117401105	30.12009652	-0.196057649
1991	32.30947811	-0.096252653	27.07754161	-0.174838377
1992	29.67623554	-0.138203225	24.97219202	-0.216825463
1993	27.1533492	-0.16707776	22.94691801	-0.245681031
1994	23.27915872	-0.197729595	19.620067	-0.275996969
…	…	…	…	…
2015	-19.19748247	-0.751181631	-16.14786961	-0.824802112
2016	-20.60952874	-0.686199375	-17.36180653	-0.759698948
2017	-21.4423159	0	-18.04057589	0
2018	-22.49845921	0	-18.9134046	0

p-ISSN 2798-253X
e-ISSN 2798-3153

2019	-23.63479673	0	-19.88749318	0

The data in Table 2 and Table 3 can be transformed into graphs, as presented in Fig. 2 and Fig. 3.

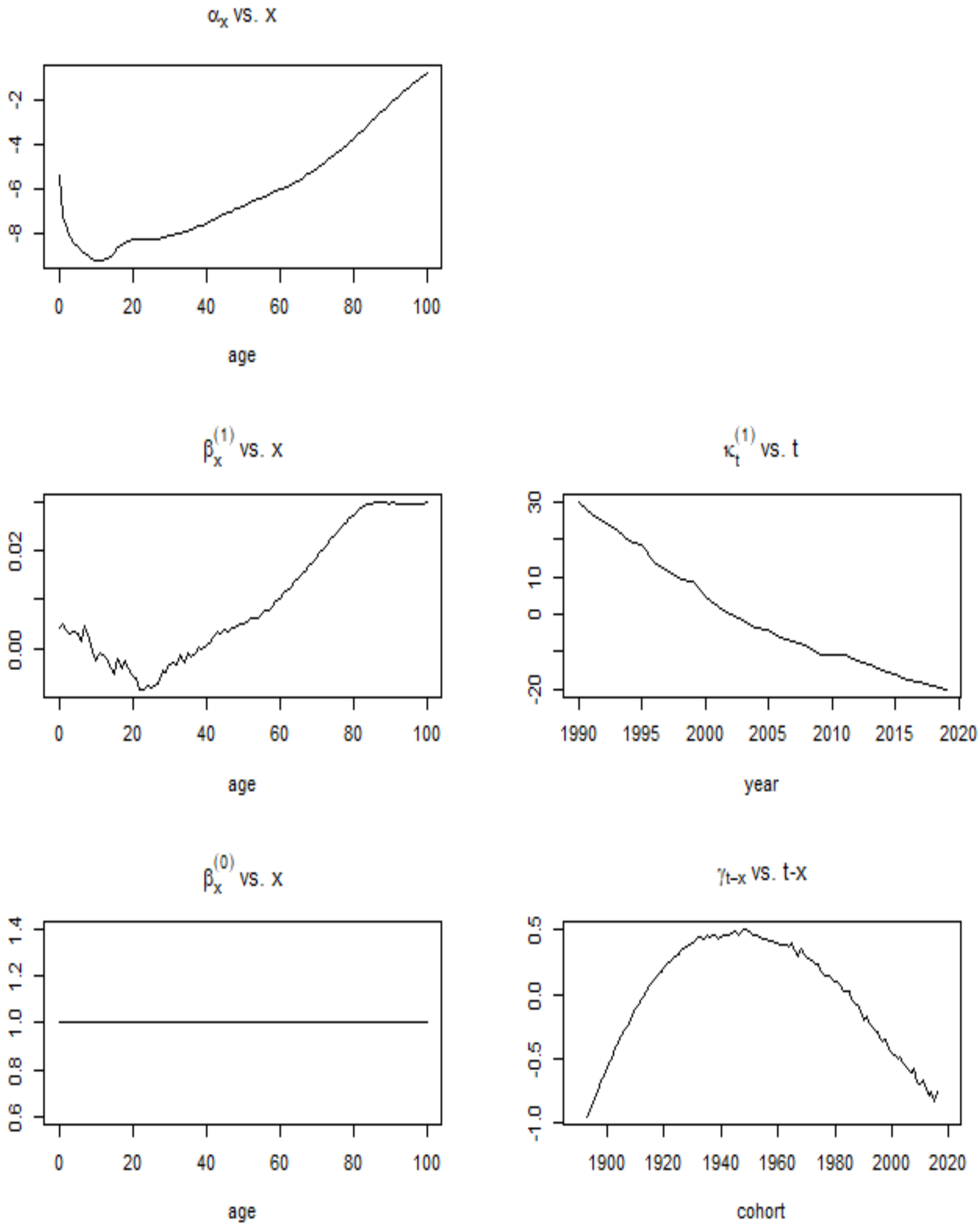

Fig. 2 Estimation results of female parameters.

p-ISSN 2798-253X
e-ISSN 2798-3153

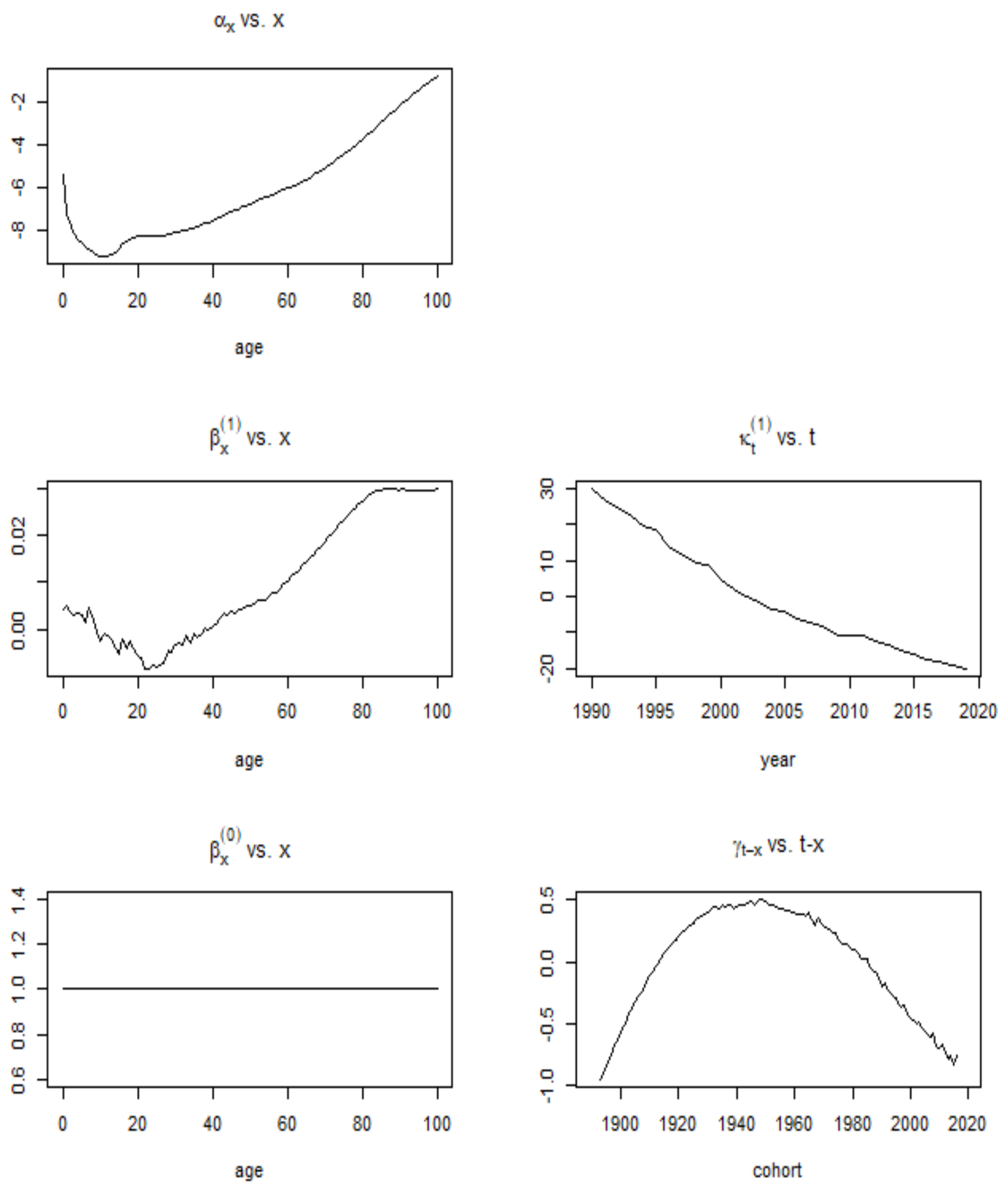

Fig. 3 Result of male parameter estimation.

Based on Table 2, Fig. 2, and Fig. 3 above, it can be seen that the average mortality rate (α_x) indicates that the death rate at the age of zero (toddlers) is quite high. At the age of toddlers and adolescents, the mortality rate has decreased. For teenagers to adults, the death rate increases, and in the elderly, the death rate is high. This means that the average death rate in adults to the elderly has a higher risk of death.

From the results of Table 1, Fig. 2, and Fig. 3, the parameters $\left(\beta_x^{(1)}\right)$describe the pattern of changes in mortality rates based on age from changes in values k_t and in Fig. 2 and Fig. 3, it can be seen that the relative speed of change occurs at the age of toddlers to adolescents. in adulthood to the elderly is relatively constant in both females and males. In Table 1, the parameter $\beta_x^{(0)}$ is constant since based on equation (7) it is determined that $\beta_x^{(0)} = 1$. Therefore, the change in the death rate based on age from the change in value ι_{t-x} is considered constant.

p-ISSN 2798-253X
e-ISSN 2798-3153

The parameter k_t describes the effect of the period in year t. In Table 3, it can be seen that the estimation k_t results are decreasing year by year, more detailed pictures can be seen in Fig. 2 and Fig. 3. The depiction of the period effect between females and males has a fairly small difference. The parameter ι_{t-x} describes the cohort effect in year t. From Table 3, it can be seen that from 1990 - 2019 the cohort effect tends to decrease which can be seen also in Fig. 2 and Fig. 3. It means that the characteristics of the data from 2019 - 2020 are different from the existing empirical data.

From the table and figure above, it is shown that in general the estimation results obtained with Generalized Lee-Carter are almost close to the actual value. To find out how well the model is for estimating its parameters, RMSE the RMSE results for each were used: 0.016710 for females and 0.016292 for males. Based on the RMSE results, the Generalized Lee-Carter model can be used to predict mortality rates.

To forecast the future, using the estimation results of the parameter k_t and ι_{t-x} obtained by using ARIMA. The time series data used is period effect and cohort effect data, namely k_t and ι_{t-x}.

To predict the death rate in the next few years, time series data from the estimation results are used and ARIMA is used. The results obtained using ARIMA show that the best model for k_t is ARIMA (0.2,2) for females and males. The AIC value is 83,25742 for females and for males with AIC 79,55025. Best model of ι_{t-x} is ARIMA (0,1,0) for females with AIC value of -32.18762 and ARIMA (1,0,0) for males with AIC value of -24.73325. Therefore, the forecast for the death rate of the Japanese population for the next few years is presented in Fig. 4.

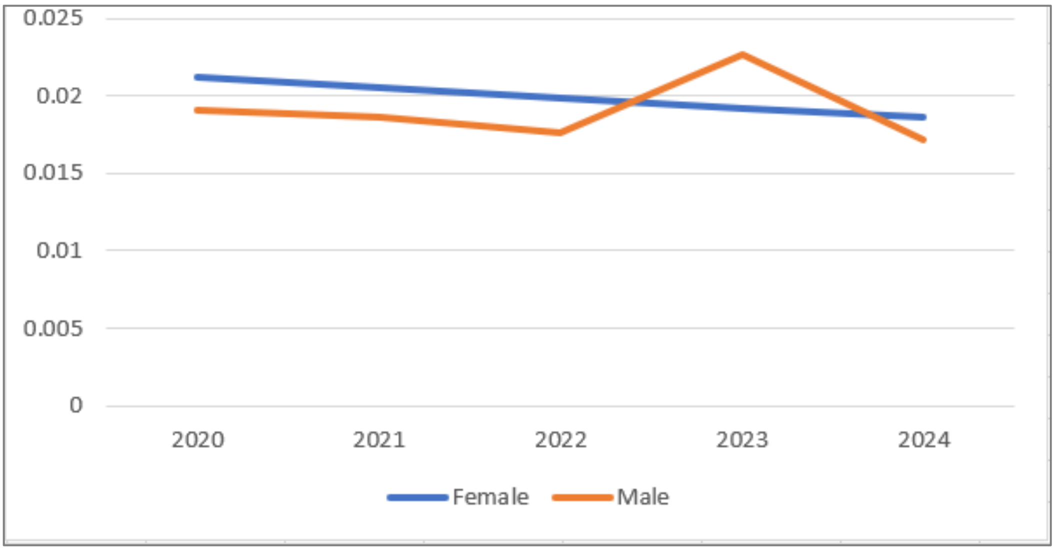

Fig. 4 Average mortality rate in the next years.

3.2. Discussion

Based on the description above, it is known that Japan, which is a developed country, has a fairly high mortality rate and holds the third rank as the country with the highest death rate. Many factors affect the high mortality rate of the Japanese population, such as the economy, depression, employment, disease, and many others. In addition to this, the low birth rate also affects, with low births, the population rate in the elderly will be higher. As shown in Fig. 2 and Fig. 3, the parameters that describe the average death rate show that from adolescence to adulthood, the mortality rate is quite high. Adolescence and adulthood are the ages where there is a productive period in doing work, the factors described previously are many in that age. In the Lee-Carter model these factors have not been captured properly, so that the generalized Lee-Carter model, with the addition of this cohort effect, provides a better picture in forecasting. Changes in the cohort effect occur due to differences

p-ISSN 2798-253X
e-ISSN 2798-3153

in generations and lifestyles in each period. This change in pattern has an impact on the death rate experienced by humans.

4. Conclusion

From the results of the discussion and study conducted by the Generalized Lee-Carter model as a general form of the Lee-Carter model in forecasting, it gives good results. In a case study conducted using Japanese population data, it gave good results. The death rate of the Japanese population as a result of forecasting the death rate in the next few periods shows that mortality *male* rate is higher than that of *female.*

In this study, there are still many shortcomings that need to be developed, especially in the study of the model so that it can be studied better. In addition to this, the Generalized Lee-Carter model can still be further developed and can be used to estimate mortality rates in other countries.

Acknowledgment

We sincerely thank Mr. I Gusti Nyoman Yudi Hartawan and I Gusti Ngurah Pujawan who have helped and guided us in completing this article so we could achieve the best results; thank you also to my teammate Kadek Jayanta for assisting in making this article well; and thank you to the Asia Student Paper Competition 2022 for providing us with the opportunity and experience.

References

[1] M.A.F. Alfana, M. Iffani, and W.A.N.P Hanif, "Mortalitas di Indonesia (Sejarah Masa Lalu dan Proyeksi ke Depan)," 05-Sep-2017. [Online]. Available: osf.io/2d7yk.

[2] D.A. Swanson and J.S Siegel, *The Methods and Materials of Demography*, 2nd ed. Bingley, U.K.: Emerald Publishing, 2004.

[3] A. Widiandari, "Fenomena Shoushika Di Jepang : Perubahan Konsep Anak," *Izumi*, vol. 5, no. 1, pp. 32–39, 2016, doi:10.14710/izumi.5.1.32-39.

[4] S.I. Trahutami, "Kasus Bunuh Diri pada Masyarakat Jepang (Penyebab dan Tren Dewasa Ini)," *Kiryoku*, vol. 1, no. 1, pp. 12–17, 2017, doi: https://doi.org/10.14710/kiryoku.v1i1.%p.

[5] M.T.B. Karo, I. Hikmatullah, M.D. Puteri, Q.N. Aulia, and N.S. Shafira, "Fenomena Shoushika: Analisis Kebijakan Pemerintah Jepang Pada Era Kepemimpinan Shinzo Abe," *TransBORDERS: Internasional Relation Journal*, vol. 4, no. 2, pp. 96–110, 2021, doi: https://doi.org/10.23969/transborders.v4i2.3939.

[6] I.G. Nyoman, Y. Hartawan, and K.M. Pranata, "Peramalan Tingkat Kematian Penduduk Kabupaten Buleleng Menggunakan Model Cairns-Blake-Dowd (CBD)," *Undiksha Press. Seminar Nasional Riset Inovatif*, 2020, pp. 36–43.

[7] R.D. Lee and L.R. Carter, "Modeling and Forecasting U.S. Mortality," *Journal of the American Statistical Association*, vol. 87, no. 419, pp. 659-671, 1992, doi:10.2307/2290201.

[8] A. Hunt and A.M. Villegas, "Robustness and Convergence in the Lee-Carter Model with Cohort Effects," *Insurance: Mathematics and Economics*, vol. 64, pp. 186–202, 2015, doi: 10.1016/j.insmatheco.2015.05.004.

[9] S. Tuljapurkar, N. Li, and C. Boe, "A Universal Pattern of Mortality Decline in the G7 Countries," *Nature*, vol. 405, pp. 789–792, 2000, doi:10.1038/35015561.

[10] Robinson *et al*., "City Research Online City, University of London Institutional Repository," *City Res. Online*, vol. 37, no. 9, pp. 1591–1601, 2008.

[11] A. Renshaw and S. Haberman, "Lee-Carter Mortality Forecasting: A Parallel Generalized Linear Modeling Approach for England and Wales Mortality Projections," *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, vol. 52, no. 1, pp. 119–137, 2003, doi:10.1111/1467-987600393.

[12] A.E. Renshaw and S. Haberman, "A Cohort-Based Extension to the Lee-Carter Model for Mortality Reduction Factors," *Insurance: Mathematics and Economics*, vol. 38, no. 3, pp. 556–570, 2006, doi:10.1016/j.insmatheco.2005.12.001.

[13] Z. Butt, S. Haberman, and H.L. Shang, "Lee-Carter Mortality Models using Iterative Fitting Algorithms," Package 'ilc' Version 1.0, Nov. 19, 2014. [Online]. Available http://cran.nexr.com/web/packages/ilc/ilc.pdf.

 p-ISSN 2798-253X
e-ISSN 2798-3153

[14] A.M. Villegas, P. Millossovich, and V.K. Kaishev, “StMoMo: Stochastic Mortality Modeling in R,” *Journal of Statistical Software*, vol. 84, no. 3, 2018, doi:10.18637/jss.v084.i03.

[15] S. Haberman and A. Renshaw, “A Comparative Study of Parametric Mortality Projection Models,” *Insurance: Mathematics and Economics*, vol. 48, no. 1, pp. 35–55, 2011, doi:10.1016/j.insmatheco.2010.09.003.

 p-ISSN 2798-253X
e-ISSN 2798-3153