Available online at http://scik.org
Commun. Math. Biol. Neurosci. 2023, 2023:115
https://doi.org/10.28919/cmbn/8184
ISSN: 2052-2541

FORECASTING INDONESIA MORTALITY RATE USING BETA AUTOREGRESSIVE MOVING AVERAGE MODEL

MUHAMMAD FAIZ AMIR ATHTHUFAIL, SINDY DEVILA*, FEVI NOVKANIZA

Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Indonesia, Depok 16424, Indonesia

Abstract: The mortality rate serves as one measure of the health sector as well as a tool for identifying populations that should receive specific health and development programs. The mortality rate can be used to determine a nation's level of welfare and standard of living. The mortality rate also affects the pricing of insurance premiums, the calculation of the benefit reserve for annuity products, actuarial risk management, and pension plans. A model is required to predict the mortality rate in the future because it is a random variable that varies over time and is in the range of (0,1). The Beta Autoregressive Moving Average (βARMA) model is a development of Beta regression and can be used to model and forecast mortality rates. Based on data on Indonesia's annual death rates from 1960 to 2020, we constructed a βARMA model for forecasting Indonesia's mortality rate. The best βARMA model was selected using Akaike's Information Criterion (AIC) value, and forecasting accuracy was assessed using Root Mean Square Error (RMSE). For Indonesia's annual mortality rate data, the best βARMA model produces an RMSE value of 0.0001.

Keywords: Akaike’s information criterion; beta regression; mortality rate; proportion; root mean square error.

2020 AMS Subject Classification: 60G10, 62M10.

---

*Corresponding author
E-mail address: sindy@sci.ui.ac.id
Received August 18, 2023

1. INTRODUCTION

The mortality rate quantifies the number of deaths per certain size of the population exposed to the death risk [1]. In various fields, such as health, epidemiology, or even national planning, mortality rate becomes an important thing to consider. In the health sector, mortality rate plays role as an indicator of health development. In other words, the lower mortality rate in a certain area implies that the health condition of that area is already good. In epidemiology, the mortality rate reflects the changing disease pattern of a country thus helping in better utilization of available resources [2]. In national planning, the mortality rate could help to identify target groups for special health and development programs [3]. Mortality rate also represents the level of welfare and quality of life in a country. Mortality rate also plays role in calculating premium prices (pricing) and calculating benefit reserves (valuation) for insurance policies and annuity products, as well as play role in actuarial risk management and pension plans.

Considering the important roles of mortality rate in various aspects, an appropriate model is needed to be able to forecast the mortality rate well. Mortality rate is a random variable that changes from time to time [4]. Therefore, previous research used a time series model to model and forecast mortality rates. One of the first models that have been used to forecast mortality rates is Lee-Carter model [5] which is introduced in 1992. However, over time, other time series models were also used to forecast mortality rates. For example, a previous study used Seasonal Autoregressive Integrated Moving Average (SARIMA) model to forecast mortality rate caused by traffic accidents in China [6]. A previous study also compares three time series models which are Autoregressive Integrated Moving Average (ARIMA), Beta Autoregressive Moving Average (βARMA), and Kumaraswamy Autoregressive Moving Average (KARMA) to forecast Brazilian mortality rates due to occupational accidents [7].

Several previous studies that discuss mortality rate modelling, usually modeled the number of deaths that occur from one period to another. Whereas mortality rate is a proportion of the number of deaths during a specified period and the population at risk of dying during that period [8]. This means, the increasing or decreasing number of deaths from one period to another, does not necessarily indicate the increasing or decreasing mortality rate. This is because mortality rate also depends on the population at risk of dying which changes from time to time. Therefore, in this study we modeled and forecasted mortality rate directly instead of the number of deaths from one period to another. Different from the number of deaths, mortality rate is a proportion, which has a characteristic of values that fall in the interval (0,1). Thus, such a model to accommodate this characteristic is required to forecast mortality rate. In the previous study, one of the time series models that was used to model and forecast mortality rate is Beta Autoregressive Moving Average (βARMA) model [9, 10]. This model was introduced by Rocha & Cribari-Neto in 2009 to be able to model a continuous random variable that has values in the interval (0,1) and observed from time to time, such as proportion. For those reasons, this study will try to model and forecast Indonesia's mortality rate using the βARMA model.

2. Beta Autoregressive Moving Average Model

The Beta Autoregressive Moving Average (βARMA) model was introduced by Rocha & Cribari-Neto in 2009. This model is a time series model that can be used to model and forecast random variables that assume values in the interval (0,1), such as proportion, and the model is based on the Beta Regression model [11] which was introduced by Ferrari & Cribari-Neto in 2004. The Beta Regression model itself is a Generalized Linear Model (GLM).

2.1. Generalized Linear Model

Generalized Linear Model (GLM) extends standard linear regression models to encompass non-normal response distributions and possibly nonlinear functions of the mean [12]. GLM has three components:

a. Random Component which specifies the response variable y and its probability distribution. The observations $y_1, \dots, y_n$ are treated as independent.
b. Linear predictor or systematic component, η_t, is a linear combination of explanatory variables $x_1, \dots, x_k$ and linear parameter $\beta_1, \dots, \beta_k$ which is shown by equation (1).

$$\eta_i = \sum_{j=1}^{k} \beta_j x_{ij}, \quad i = 1, \dots, n \tag{1}$$

c. Link function, which is a function $g(\cdot)$ that relates the linear predictor to the mean of the random component ($\eta_i = g(\mu_i)$) where $\mu_i = E(y_i),\ i = 1, \dots, n.$ $g(\cdot)$ is a monotonic and differentiable function. Thus $g(\cdot)$ links μ_i to explanatory variables through the formula:

$$g(\mu_i) = \sum_{j=1}^{k} \beta_j x_{ij}\,,\ \ i = 1, \dots, n \tag{2}$$

2.2. Beta Regression

Beta Regression model was introduced by [11] and [13] with the purpose to be able to model a special situation where the response variable (y) measured continuously in an interval unit, such as $0 < y < 1$. This model became the solution to the problem faced by linear regression model in modeling this special situation, because fitted values that is produced for the response variable could be exceeding the upper bound and/or lower bound.

The Beta Regression assumes that the response variable, y, follows a beta distribution with the parameters a and b. The beta distribution is very flexible in modeling proportion because it has a lot of shapes that depend on the values of its parameters. The density function beta is shown by the following formula:

$$f(y; a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} y^{a-1}(1-y)^{b-1}, \quad 0 < y < 1, \tag{3}$$

where $a > 0$, $b > 0$ and $\Gamma(\cdot)$ is a gamma function. The mean and variance of y are shown by the following equation,

$$E(y) = \frac{a}{a+b} \tag{4}$$

$$Var(y) = \frac{ab}{(a+b)^2(a+b+1)} \tag{5}$$

Then [11] defines a different parameterization of beta density so it contains a precision parameter. Let $\mu = \frac{a}{a+b}$ and $\varphi = a + b$, we can rewrite the mean and variance of the response variable, y, as follows,

$$E(y) = \mu \tag{6}$$

$$Var(y) = \frac{V(\mu)}{1+\varphi} \tag{7}$$

where $V(\mu) = \mu(1-\mu)$, μ is the mean of the response variable and φ is the precision parameter, which means for a certain value of μ, the larger the value of φ, the smaller the variance of y. Thus, we can rewrite the beta density using the new parameter,

$$f(y;\mu,\varphi) = \frac{\Gamma(\varphi)}{\Gamma(\mu\varphi)\Gamma\big((1-\mu)\varphi\big)} y^{\mu\varphi-1}(1-y)^{(1-\mu)\varphi-1}, \quad 0 < y < 1, \tag{8}$$

where $0 < \mu < 1$ and $\varphi > 0$.

If $y_1, \dots, y_n$ are independent random variables, where every y_t, $t = 1, \dots, n$ follows the beta density as shown in equation (8) with μ_t as the mean and φ is an unknown precision parameter, the Beta Regression model can be obtained by assuming that the function of mean y_t can be written as the following formula,

$$g(\mu_t) = \sum_{j=1}^{k} \beta_j x_{tj} = \eta_t \tag{9}$$

where $\boldsymbol{\beta} = (\beta_1, \dots, \beta_k)^T$ is a vector of unknown regression parameters and $x_{t1}, \dots, x_{tk}$ are the k-th observation covariate where $k < n$ which is assumed to be known and fixed. Whereas $g(\cdot)$ is the link function, a monotonic and twice differentiable function that maps (0,1) to $\mathbb{R}$. According to [14], several link functions that can be used are:

a. Logit link function

$$g(\mu) = \log\left(\frac{\mu}{1-\mu}\right)$$

b. Probit link function

$$g(\mu) = \Phi^{-1}(\mu)$$

where $\Phi(\cdot)$ is the standard normal distribution function.

c. Complementary log-log link function

$$g(\mu) = \log\{-log(1-\mu)\}$$

d. Log-log link function

$$g(\mu) = -\log\{-\log(\mu)\}$$

2.3. Autoregressive Moving Average

Autoregressive Moving Average (ARMA) process is a combination of the Autoregressive (AR) and Moving Average (MA) processes. According to [15], Autoregressive processes are regression on themselves (past values of the response variable). Whereas Moving Average processes are a special case of a general linear process, which is a weighted and finite linear combination of white noises (e_t) where the weight is not zero.

An Autoregressive process $\{Y_t\}$ with order p, AR (p), is shown by the following equation,

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t \tag{10}$$

where e_t is white noise term that is assumed to be independent from $Y_{t-1}, \dots, Y_{t-p}$ and $e_t \sim N(0, \sigma^2)$. On the other side, a Moving Average process with order q, MA (q), is shown by the following equation,

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \tag{11}$$

where e_t are white noise terms and $\theta_i \neq 0$. Therefore, An Autoregressive Moving Average (ARMA (p, q)) which is a combination of Autoregressive process and Moving Average process can be shown by the following equation,

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e \tag{12}$$

2.4. Beta Autoregressive Moving Average

[8] and [9] introduced a time series model based on the class of Beta Regression models of [11]. The model is called Beta Autoregressive Moving Average (βARMA) which can be used to model and forecast response variables that assume values in the interval (0,1), such as proportion.

Let y_t, $t = 1, \dots, n$, are continuous random variables at period t with the values in the interval (0,1). Assume the conditional distribution of each y_t, given the previous information set $\mathcal{F}_{t-1}$ (the smallest σ-algebra such the variables $y_1, \dots, y_{t-1}$ are measurable), follows the beta distribution with the conditional density of y_t given $\mathcal{F}_{t-1}$ is shown by the following equation,

$$f(y_t|\mathcal{F}_{t-1}) = \frac{\Gamma(\varphi)}{\Gamma(\mu_t\varphi)\Gamma\big((1-\mu_t)\varphi\big)} y_t^{\mu_t\varphi - 1}(1-y_t)^{(1-\mu_t)\varphi - 1}, \quad 0 < y_t < 1, \tag{13}$$

where the conditional mean is $E(y_t|\mathcal{F}_{t-1}) = \mu_t$ and the conditional variance is $Var(y_t|\mathcal{F}_{t-1}) = \frac{V(\mu_t)}{(1+\varphi)}$, with $V(\mu_t) = \mu_t(1-\mu_t)$. We also assume that the covariates $\boldsymbol{x}_t$, $t = 1, \dots, n$, where $\boldsymbol{x}_t = (x_{t1}, \dots, x_{tk})^T$ are non-random with $k < n$.

In the βARMA model, the systematic component consists not only the linear predictor but there is also an additional component, τ_t, which allows Autoregressive Moving Average term to be included additively. Therefore, a general model for μ_t is given by the following equation,

$$g(\mu_t) = \eta_t = \boldsymbol{x}_t^T\boldsymbol{\beta} + \tau_t \tag{14}$$

where $\boldsymbol{\beta} = (\beta_1, \dots, \beta_k)^T$ is the unknown linear parameter.

To show the representation of τ_t as ARMA component, we need to define an ARMA (p, q) model initially as function of a term ξ_t, such that $\xi_t = g(y_t) - \boldsymbol{x}_t^T\boldsymbol{\beta}$, where:

$$\xi_t = \alpha + \sum_{i=1}^{p} \phi_i \xi_{t-i} - \sum_{j=1}^{q} \theta_j r_{t-j} + r_t \tag{15}$$

r_t denotes a random error which is defined as $r_t = g(\mu_t) - \eta_t$ and $\alpha \in \mathbb{R}$ is a constant. Then we assumed $E(r_t|\mathcal{F}_{t-1}) = 0$. Note that because $\xi_{t-i} \in \mathcal{F}_{t-1}$, $i > 0$ and $E(\xi_t|\mathcal{F}_{t-1}) \approx \tau_t$, by taking conditional expectations with respect to the $\mathcal{F}_{t-1}$ in (15), we obtained the approximate model which is shown by,

$$\tau_t = \alpha + \sum_{i=1}^{p} \phi_i \xi_{t-i} - \sum_{j=1}^{q} \theta_j r_{t-j} \tag{16}$$

Because $\xi_t = g(y_t) - \boldsymbol{x}_t^T\boldsymbol{\beta}$, then equation (16) can be rewrite,

$$\tau_t = \alpha + \sum_{i=1}^{p} \phi_i \{g(y_{t-i}) - \boldsymbol{x}_{t-i}^T\boldsymbol{\beta}\} - \sum_{j=1}^{q} \theta_j r_{t-j} \tag{17}$$

where $\boldsymbol{x}_t \in \mathbb{R}^k$, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_k)^T$, $k < n$, $p, q \in \mathbb{N}$, with p as the Autoregressive order, q as the Moving Average order, ϕ as the Autoregressive parameter, θ as the Moving Average and r_t is a random error. Since the $\tau_t = g(\mu_t) - \boldsymbol{x}_t^T\boldsymbol{\beta}$, the βARMA general model for the mean μ_t is,

$$g(\mu_t) = \alpha + \boldsymbol{x}_t^T\boldsymbol{\beta} + \sum_{i=1}^{p} \phi_i \{g(y_{t-i}) - \boldsymbol{x}_{t-i}^T\boldsymbol{\beta}\} - \sum_{j=1}^{q} \theta_j r_{t-j} \tag{18}$$

The equation (18) is then called βARMA (p, q). Fitted values and forecasted values that are produced by the βARMA model will have the values in the interval (0,1). Parameter estimation can be carried out using the Conditional Maximum Likelihood Estimation (see, [15, 16, 17, 18]). But most of the time the solution of the system of equations does not have a closed form.

Hence, it has to be numerically obtained by maximizing the conditional log-likelihood function using a nonlinear optimization algorithm, such as a quasi-Newton algorithm [19] like Broyden, Fletcher, Goldfarb, and Shanno (BFGS) method [20]. Figure 1 shows the flowchart of the construction of the βARMA model.

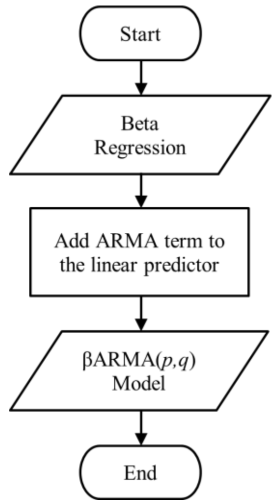

FIGURE 1. Flowchart of the construction of the βARMA model

2.5. Forecasting using the βARMA Model

In order to forecast using the βARMA model, we need to find the appropriate βARMA model for the given information. In other words, we need to find the order of the βARMA model that will produce the best model for the given dataset. After obtaining the best βARMA model, we then can perform the forecasting to obtain forecasted values for the next several periods. Therefore, in this study we follow some steps to implement the βARMA model to produce the forecast. Figure 2 shows the steps in terms of a flowchart.

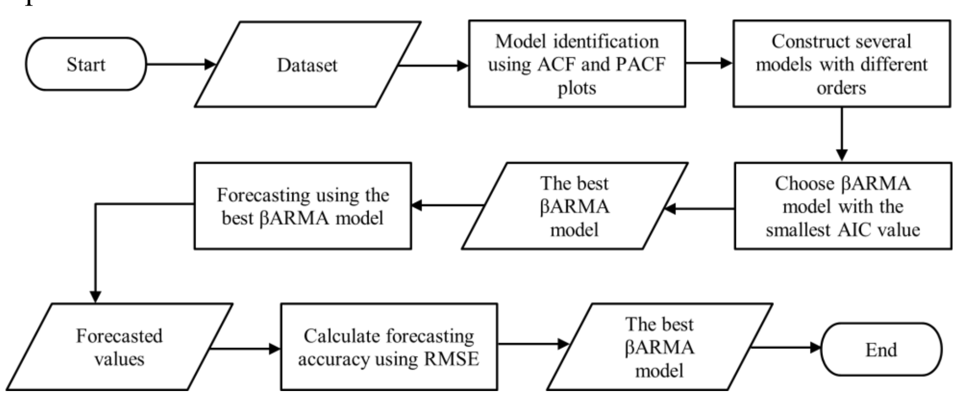

FIGURE 2. Flowchart of implementing the βARMA model

The first step is model identification. In this step, we perform model identification by analyzing the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots. This will help us to determine the order of the βARMA model. Even though the orders obtained from this step will not always give us the best model, it gives us some clues about the appropriate orders.

The second step is to construct several βARMA models with different orders. We estimate the parameters of each model using the Conditional Maximum Likelihood Estimation method. After we obtain the estimated parameters of each model, we then can proceed to the next step, which is model selection.

In the model selection step, we performed the model selection by choosing the model that produces the smallest Akaike's Information Criterion (AIC) value [21, 22]. The chosen model will be our best βARMA model. We will use the best βARMA to perform forecasting for the next several periods.

The next step is forecasting. Forecasting for the h-step ahead, $h = 1, 2, \ldots$, can be computed as

$$\hat{\mu}_{n+h} = g^{-1}\left(\hat{\alpha} + \boldsymbol{x}_{n+h}^{T}\widehat{\boldsymbol{\beta}} + \sum_{i=1}^{p}\hat{\phi}_i\left\{g(y_{n+h-i}) - \boldsymbol{x}_{n+h-i}^{T}\widehat{\boldsymbol{\beta}}\right\} + \sum_{j=1}^{q}\hat{\theta}_j\hat{r}_{n+h-j}\right) \tag{19}$$

where,

$$g(y_t) = \begin{cases} g(\widehat{\mu_t}), & jika\ t > n \\ g(y_t), & jika\ t \leq n \end{cases} \text{ and } \widehat{r_t} = \begin{cases} 0, & jika\ t > \\ g(y_t) - g(\widehat{\mu_t}), & jika\ t \leq \end{cases}$$

$\hat{\alpha}, \widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{\phi}}, \widehat{\boldsymbol{\theta}}, \hat{\varphi}$ respectively are the Conditional Maximum Likelihood Estimator of $\alpha, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\theta}, \varphi$, $\hat{\mu}_t$ is the estimated value of μ_t, and $\widehat{r_t}$ is the estimated value of r_t.

Finally, after obtaining the forecasted value for the next several periods, we then can calculate the forecasting accuracy. In this study, we calculate the forecasting accuracy using the Root Mean Square Error (RMSE) criterion [23, 24]. The best way to choose the best forecasting model is by finding a model with the smallest RMSE [25].

3. MAIN RESULTS

3.1. Data

In this study, we used Indonesia's annual mortality rate data (Crude Death Rate) from the year 1960 through 2020 which can be obtained from https://data.worldbank.org [26]. From this data, we removed the last 6 observations which then we used as validation for the forecasting values produced by the model. We applied the Beta Autoregressive Moving Average (βARMA) model to the data and used the model to forecast the mortality rate for the next six periods. The βARMA model is a time series model developed from the Beta Regression model. The RStudio with R programming language is used to conduct the analysis.

3.2. Time Series Plot and Model Identification

Figure 3 shows the time series plot of Indonesia's annual mortality rate data from the year 1960 through 2020. From the figure, we could see that Indonesia's mortality rate tends to decrease each year. We also provided the ACF and PACF plots of Indonesia's annual mortality rate data in Fig. 4. We could see that the ACF plot shows a gradually decreasing value, whereas the PACF plot shows the value cut off after lag one. This tells us that there is an Autoregressive characteristic in Indonesia's annual mortality rate data with the order of one.

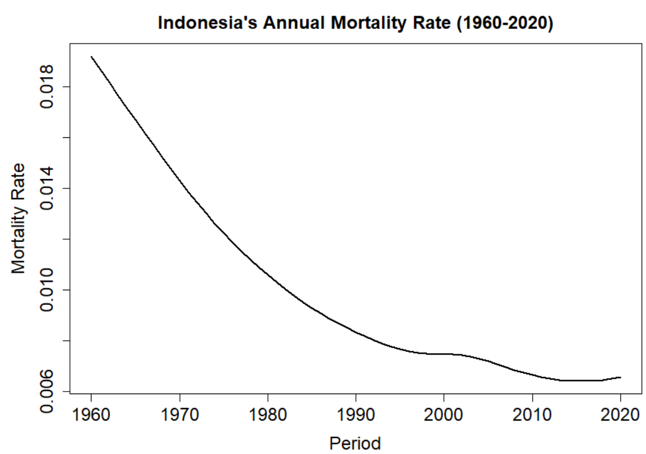

FIGURE 3. Time Series Plot of Indonesia's Annual Mortality Rate from 1960 through 2020

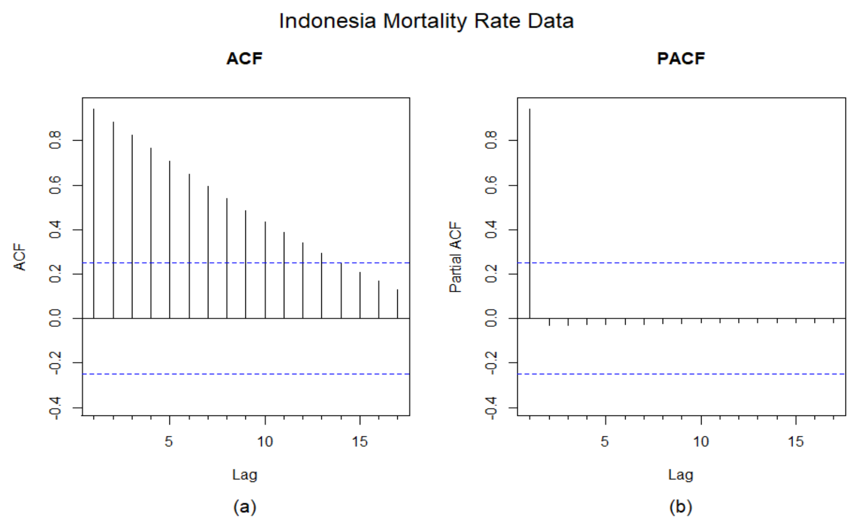

FIGURE 4. (a) ACF plot and (b) PACF plot of Indonesia's Annual Mortality Rate Data

3.3. The Best βARMA Model

To find the best βARMA model, we constructed all βARMA models with the order less than or equal to 4. This means we constructed all possible βARMA models with the order of $p \in 0, \dots, 4$ and $q \in 0, \dots, 4$. We used the βARMA model's script that was made by [27]. All of the considered βARMA models and their AIC values are shown in Table 1. After constructing the models, based on the model selection criterion, we chose the model that produced the smallest AIC value as our best βARMA model. Therefore, we obtain βARMA (4,4) as our best βARMA model.

TABLE 1. Considered βARMA models

Model	AIC	Model	AIC
βARMA(*0,1*)	-497.8532	βARMA(*2,3*)	-1057.901
βARMA(*0,2*)	-500.84	βARMA(*2,4*)	-1036.205
βARMA(*0,3*)	-502.3266	βARMA(*3,0*)	-1134.625
βARMA(*0,4*)	-494.5602	βARMA(*3,1*)	-1150.252
βARMA(*1,0*)	-904.6842	βARMA(*3,2*)	-1154.27
βARMA(*1,1*)	-939.1526	βARMA(*3,3*)	-1161.317
βARMA(*1,2*)	-936.1933	βARMA(*3,4*)	-1138.121
βARMA(*1,3*)	-923.0501	βARMA(*4,0*)	-1162.515
βARMA(*1,4*)	-906.8183	βARMA(*4,1*)	-1160.515
βARMA(*2,0*)	-1046.522	βARMA(*4,2*)	-1158.515
βARMA(*2,1*)	-1073.214	βARMA(*4,3*)	-1165.263
βARMA(*2,2*)	-1075.76	**βARMA(*4,4*)**	**-1171.602**

Table 2 shows the estimated parameters of βARMA (4,4). We could see that there are two parameters which are the parameter θ_2 and θ_3 with the $\Pr(>|z|)$ or $p-value$ greater than the significance level of 0.05, thus, these two parameters are not significant. Therefore, we considered a new βARMA model, which is βARMA (4,4) without the parameter θ_2 and θ_3. The estimated parameters of the final βARMA (4,4) model are shown in Table 3. Table 3 shows that all the $\Pr(>|z|)$ or $p-value$ of the parameters are very small, which means all parameters are significant.

TABLE 2. Estimated Parameters of βARMA (4,4) Model

Parameter	**Estimate**	**Standard Error**	z **value**	$\Pr(>\|z\|)$
α	-2.8000×10^{-3}	0.0000	70.6264	0.0000
ϕ_1	3.6479	3.5000×10^{-3}	1039.0909	0.0000
ϕ_2	-5.0986	9.9000×10^{-3}	513.3043	0.0000
ϕ_3	3.2363	9.6000×10^{-3}	336.6918	0.0000
ϕ_4	-7.8620×10^{-1}	3.2000×10^{-3}	247.2305	0.0000
θ_1	-5.4890×10^{-1}	7.2000×10^{-3}	75.9135	0.0000
θ_2	8.5000×10^{-3}	6.7000×10^{-3}	1.2737	0.2028
θ_3	-9.7000×10^{-3}	6.8000×10^{-3}	1.4122	0.1579
θ_4	-7.6080×10^{-1}	6.6000×10^{-3}	114.8237	0.0000
φ	1.2647×10^{9}	2.5044×10^{8}	5.0497	0.0000
	Log-likelihood		595.801	
	AIC		-1171.602	

TABLE 3. Estimated Parameters of Final βARMA (4,4) Model without Parameter θ_2 and θ_3

Parameter	**Estimate**	**Standard Error**	z **value**	$\Pr(>\|z\|)$
α	-2.7000×10^{-3}	0.0000	76.7530	0.0000
ϕ_1	3.6606	3.5000×10^{-3}	1043.0349	0.0000
ϕ_2	-5.1288	1.0000×10^{-2}	513.9448	0.0000
ϕ_3	3.2589	9.7000×10^{-3}	336.5534	0.0000
ϕ_4	-7.9120×10^{-1}	3.2000×10^{-3}	246.9305	0.0000
θ_1	-4.8640×10^{-1}	5.4000×10^{-3}	89.4436	0.0000
θ_4	-8.1110×10^{-1}	5.1000×10^{-3}	157.7763	0.0000
φ	1.2647×10^{9}	2.5044×10^{8}	5.0497	0.0000
	Log-likelihood		595.801	
	AIC		-1171.602	

From the analysis that has been done, we can conclude that the best βARMA model based on the smallest AIC value for Indonesia's annual mortality rate data is shown by the following equation:

$$g(\hat{\mu}) = \hat{\alpha} + \hat{\phi}_1 g(y_{t-1}) + \hat{\phi}_2 g(y_{t-2}) + \hat{\phi}_3 g(y_{t-3}) + \hat{\phi}_4 g(y_{t-4}) + \hat{\theta}_1 r_{t-1} + \hat{\theta}_4 r_{t-4} \tag{20}$$

where $(\hat{\alpha}, \hat{\phi}_1, \hat{\phi}_2, \hat{\phi}_3, \hat{\phi}_4, \hat{\theta}_1, \hat{\theta}_4)$ are the estimated parameters that are shown in Table 3. We also provided the plot of the fitted final βARMA (4,4) model for Indonesia's annual mortality rate data as shown in Fig. 5.

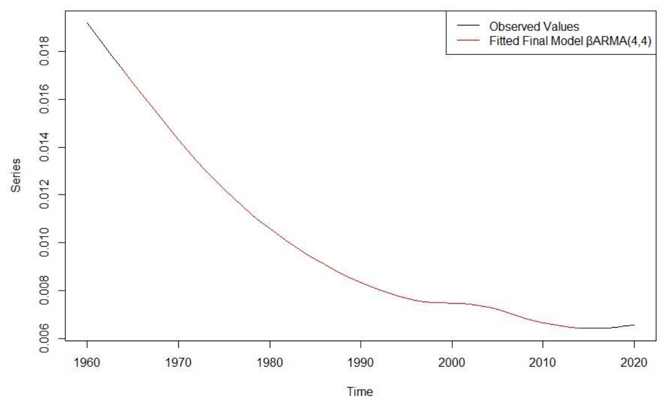

FIGURE 5. Fitted Final βARMA (4,4) model vs. Observed Data

3.4. Forecasting Indonesia's Annual Mortality Rate

Using the best βARMA model that we have obtained, we then perform the forecasting for six periods ahead. The produced forecast is then compared to the six observations that have been separated from the data before. Table 4 shows the comparison between forecasted values and observed values. We also calculated the out-of-sample forecasting accuracy using the RMSE criterion and we obtained the RMSE value of 0.0001, which is quite small. Therefore we concluded that the final βARMA (4,4) model which is the βARMA(4,4) model without parameter θ_2 and θ_3 could perform well and produce a fairly good result based on the small RMSE value. We also provided the plot of observed values and forecasted values of Indonesia's annual mortality rate which is shown in Fig.6.

TABLE 4. Observed Values and Forecasted Values of The Final βARMA (4,4) Model

Year	**Observed Values**	**Forecasted values of Final βARMA (4,4) Model**
2015	0.006419	0.006413
2016	0.006418	0.006398
2017	0.006433	0.006388
2018	0.006465	0.006381
2019	0.006510	0.006373
2020	0.006567	0.006362
	RMSE	0.000108

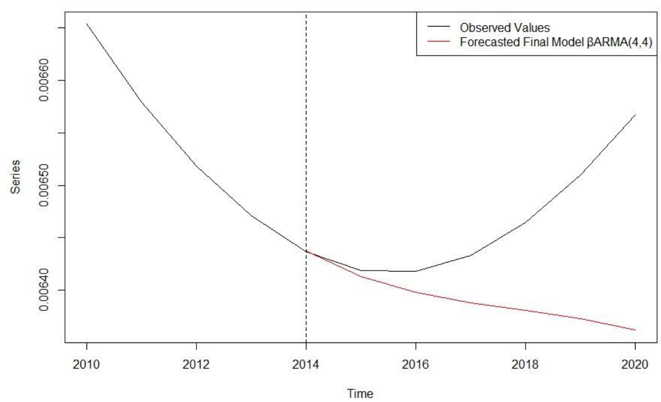

FIGURE 6. Observed Values vs. Forecasted Values of The Final βARMA (4,4) Model for Indonesia's Annual Mortality Rate From 2010 through 2020

4. CONCLUSIONS

The mortality rate is a time series that has the values in the interval (0,1), thus it requires a time series model that could represent the changes in mortality rate from time to time. βARMA(p,q) model is one of the time series models that can be used to model mortality rates. This model is based on the class Beta Regression models. Parameter estimation could be carried out using the Conditional Maximum Likelihood Estimation method because in the estimation process the conditional density given the information of previous periods is required. Model selection could be done by choosing the model that produces the smallest AIC value. The best βARMA model is then used to forecast the mortality rate in the future. The accuracy of forecasting is measured using the RMSE criterion.

ACKNOWLEDGMENTS

This research is funded by Hibah Riset Penugasan FMIPA UI
(Grant No. 016/UN2.F3.D/PPM.00.02/2022).

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interests.

REFERENCES

[1] K. Syuhada, A. Hakim, Stochastic modeling of mortality rates and Mortality-at-Risk forecast by taking conditional heteroscedasticity effect into account, Heliyon. 7 (2021), e08083. https://doi.org/10.1016/j.heliyon.2021.e08083.

[2] G. Gururaj, Importance of mortality statistics in epidemiology of neurological disorders, NIMHANS J. 10 (1992), 17-21.

[3] United Nations, Levels and trends of mortality since 1950, Study 74. United Nations, Department of International Economic and Social Affairs, New York, 1982.

[4] D. Mitchell, P. Brockett, R. Mendoza-Arriaga, et al. Modeling and forecasting mortality rates, Insurance: Math. Econ. 52 (2013), 275-285. https://doi.org/10.1016/j.insmatheco.2013.01.002.

[5] R.D. Lee, L.R. Carter, Modeling and forecasting U.S. mortality, J. Amer. Stat. Assoc. 87 (1992), 659-671. https://doi.org/10.1080/01621459.1992.10475265.

[6] X. Zhang, Y. Pang, M. Cui, L. Stallones, H. Xiang, Forecasting mortality of road traffic injuries in China using seasonal autoregressive integrated moving average model, Ann. Epidemiol. 25 (2015), 101-106. https://doi.org/10.1016/j.annepidem.2014.10.015.

[7] C. Melchior, R.R. Zanini, R.R. Guerra, D.A. Rockenbach, Forecasting Brazilian mortality rates due to occupational accidents using autoregressive moving average approaches, Int. J. Forecast. 37 (2021), 825-837. https://doi.org/10.1016/j.ijforecast.2020.09.010.

[8] M. Porta, A dictionary of epidemiology, 6th ed., Oxford University Press, New York, 2014.

[9] A.V. Rocha, F. Cribari-Neto, Beta autoregressive moving average models, TEST. 18 (2008), 529-545. https://doi.org/10.1007/s11749-008-0112-z.

[10] A.V. Rocha, F. Cribari-Neto, Erratum to: Beta autoregressive moving average models, TEST. 26 (2017), 451-459. https://doi.org/10.1007/s11749-017-0528-4.

[11] S. Ferrari, F. Cribari-Neto, Beta regression for modelling rates and proportions, J. Appl. Stat. 31 (2004), 799-815. https://doi.org/10.1080/0266476042000214501.

[12] A. Agresti, Foundations of linear and generalized linear models, Wiley, Hoboken, 2015.

[13] R. Kieschnick, B.D. McCullough, Regression analysis of variates observed on (0, 1): percentages, proportions and fractions, Stat. Model. 3 (2003), 193-213. https://doi.org/10.1191/1471082x03st053oa.

[14] F. Cribari-Neto, A. Zeileis, Beta regression in R, J. Stat. Softw. 34 (2010), 1-24. https://doi.org/10.18637/jss.v034.i02.

[15] J.D. Cryer, K.S. Chan, Times series analysis with applications in R, 2nd Ed., Springer, New York, 2008.

[16] R.A. Fisher, On the mathematical foundations of theoretical statistics, Phil. Trans. R. Soc. Lond. A. 222 (1922), 309-368. https://doi.org/10.1098/rsta.1922.0009.

[17] R.V. Hogg, J.W. McKean, A.T. Craig, Introduction to mathematical statistics, 7th ed., Pearson Education, 2013.

[18] C. Lee, Maximum Likelihood Estimation, Chapter 17, Institute of Economics, National Sun Yat-Sen University, Taiwan, 2019.

[19] W.C. Davidon, Variable metric method for minimization, Office of Scientific and Technical Information (OSTI), 1959. https://doi.org/10.2172/4252678.

[20] J. Nocedal, S.J. Wright, Numerical optimization, 2nd ed. New York, 2006.

[21] H. Akaike, Information theory and an extension of the maximum likelihood principle, Akadémiai Kiadó, Budapest, Hungary, 1973.

[22] H. Akaike, A new look at the statistical model identification, IEEE Trans. Automat. Contr. 19 (1974), 716-723. https://doi.org/10.1109/tac.1974.1100705.

[23] S.G. Makridakis, S.C. Wheelwright, R.J. Hyndman, Forecasting: methods and application, 3rd ed. Wiley, 1997.

[24] R.J. Hyndman, A.B. Koehler, Another look at measures of forecast accuracy, Int. J. Forecast. 22 (2006), 679-688.

https://doi.org/10.1016/j.ijforecast.2006.03.001.

[25] R.J. Hyndman, G. Athanasopoulos, Forecasting: principles and practice, Monash University, Australia, 2018. OTexts.com/fpp2.

[26] The World Bank, Death rate, crude (per 1,000 people) - Indonesia, 2020. https://data.worldbank.org/indicator/SP.DYN.CDRT.IN?locations=ID.

[27] V.T. Scher, F. Cribari-Neto, G. Pumi, F.M. Bayer, Goodness-of-fit tests for βARMA hydrological time series modeling, Environmetrics. 31 (2019), e2607. https://doi.org/10.1002/env.2607.