Pyramid Population Prediction using Age Structure Model

Heni Widayani[1], Nuning Nuraini[2], Anita Triska[3]

[1,2,3]Department of Mathematics, Faculty of Mathematics and Natural Sciences,
Institut Teknologi Bandung

Email: heni.w@s.itb.ac.id, nuning@math.itb.ac.id, a.triska@students.itb.ac.id

ABSTRACT

Population composition in a country by sex and age-structure often illustrated through the Population Pyramid. In this study, an age-structure model will be constructed to predict the population pyramid shape in the coming year. It is assumed that changes in population are affected by natality and mortality number in each age group, ignoring migration rates. The proposed age structure model formulated as a first-order partial differential equation with the non-negative initial condition. The boundary condition is given by the number of births which is proportional to the number of women at childbearing age. Then, this age structure model implemented utilizing United Nations Data to predict population pyramids of Indonesia, Brazil, Japan, the USA, and Russia. The population pyramid prediction of the five countries shows different characteristics, according to whether it is a developing or developed country. The results of this study indicate that the age structure model can be used to predict the composition of the population in a country in the next few years. Indonesia is predicted to be the highest populated country in 2066, compared to the other four countries. This result can be used as a reference for the government to plan policies and strategies according to age groups to control population explosion in the future.

Keywords: Population Pyramid; Prediction; Age-Structure

INTRODUCTION

The study of population characteristics which we called demography can be complex. Population pyramids are a useful tool for understanding the structure and composition of populations because they graphically portray many aspects of a population, such as sex ratios and age structure [1]. Structured population models distinguish among individuals according to characteristics such as age to determine the birth and death rates. Age is one of the most important characteristics in the modeling of populations [2]. Individuals of different ages may have different reproduction and survival capacities. Therefore, it is often useful to divide a population into different groups by age, for obtaining a better description of the population dynamics.

Mathematical models of populations incorporating spatial structure, age structure, or another structuring of individuals with continuously varying properties have been developed by many researchers [3]. Nevertheless, these models are mathematically complex and their scientific applicability depends on extensive parametric inputs. Leal-Ramirez [4] proposed a discrete age-structured population model within a fuzzy cellular structure. Russell [5] shows that a fluctuating of mortality and reproductive rate gives a cumulative extinction probability analytically. Abia [6] considers the numerical solution

of age-structured population models using characteristic curves methods and finite difference methods.

Euler suggested to modeled the growth of the human population geometrically [7]. This idea was developing by Malthus [8] who predicted the exponential growth of the human population through time. This model assumed that the population depended on constant fertility and mortality rates. Denoted that the constant fertility rate as α, the constant mortality rate as μ, and the size of the human population in time as $U(t)$ then the population dynamics was determined by the *Malthus Law*

$$\frac{dU(t)}{dt} = (\alpha - \mu)U(t) \tag{1}$$

Verhulst [9] modified the Malthus model became the famous logistic model by introducing a maximum size K for the population size which can be written as

$$\frac{dU(t)}{dt} = (\alpha - \mu)\left(1 - \frac{U(t)}{K}\right)U(t) \tag{2}$$

This logistic model was applied by Lotka [10] to fit the growth curve of the United States population in the thirties. These experiments provided growth curves that agreed with the prediction of the logistic equations. All of this model cannot capture the composition age of the population, since it just a function of time. For demography purposes, we also need to know the age composition of the population as an important issue.

The importance of population age-structured prediction of a country due to the government can plan the right strategies based on age-group for the people welfare. The specific regulation with targeting a specific age group expected to be more efficient. For young people, the government can focus on improving nutrition and health so that infant or child mortality can be reduced. For adolescents, policies related to education become the primary. Socialization about the negative side of promiscuity and drugs also become a major issue. For adulthood, economic policies become the main issues because this age is a productive age in which the individual became an economic milestone in a country.

METHODS

Age modeled as a continuous variable determined by individual year of birth. These continuous variables are then categorized by the 5 yearly age group as a discrete-time interval. The 3-year age group consists of children who have just celebrated their third birthday, children who are 3.5 years old, 3.8 years old, and so on. In this age-structured model, all individuals in the same age-group are assumed to have the same death rate. The age of an individual denoted by x. Since the newborn age is 0 years and the oldest is 105 years, then $x \in [0,105]$. Individuals over 105 years will be classified into the oldest age-group.

Let $U(x,t)$ denote the number of individuals of age x living at time t in a certain domain. To model the population dynamic we have to consider an aging process, i.e.,

$$U(x, t + \delta t) = U(x - \delta t, t) \tag{3}$$

Besides the aging processes, we also take into account the death process in which the individual dies at any age with age-dependent probability, denoted by $\mu(x)$. We neglect

the immigration process in the domain since the immigration number insignificant compared to the total population of a country.

Based on aging and death process, we consider the following discrete model

$$U(x, t + \delta t) = U(x - \delta t, t) - \mu(x)U(x, t)\delta t \tag{4}$$

A Taylor expansion for $0 \ll \delta t \ll 1$ yields

$$\begin{cases} U(x, t + \delta t) = U(x, t) + \delta t \dfrac{\partial U(x, t)}{\partial t} + \dfrac{1}{2}(\delta t)^2 \dfrac{\partial U(x, t)^2}{\partial t^2} + \sigma(\delta t^3) \\ U(x, t - \delta t) = U(x, t) - \delta t \dfrac{\partial U(x, t)}{\partial t} + \dfrac{1}{2}(\delta t)^2 \dfrac{\partial U(x, t)^2}{\partial t^2} + \sigma(\delta t^3) \end{cases} \tag{5}$$

And hence we obtain in leading order

$$\frac{\partial U(x, t)}{\partial t} + \frac{\partial U(x, t)}{\partial x} = -\mu(x)U(x, t) \tag{6}$$

where $(x, t) \in \mathbb{R}_2^+$ with non-negative initial condition

$$U(t = 0, x) = U_0(x) \tag{7}$$

The $U_0(x)$ is the age structure at time $t = 0$. The boundary condition has interpretation as the number of birth in time t (population who have zero-age $x = 0$) which is proportional to the number of female people in reproductive age [11].

The number of new-born in year t_{n+1} assumed to be proportional to the women at childbearing age in year t_n. A new-born is assumed has equal probabilities to be a male or female baby. The birth rate is a function of age x and denoted by $\alpha(x)$, then we get the boundary condition for (6) is

$$U(t, x = 0) = \int_{x_0}^{x_1} \alpha(s)U(t, s)\, ds \tag{8}$$

So, the age-structured population model including birth rate can be written as

$$\begin{cases} \dfrac{\partial U(x, t)}{\partial t} + \dfrac{\partial U(x, t)}{\partial x} = -\mu(x)U(x, t) \\ U(x, 0) = U_0(x) \\ U(0, t) = \displaystyle\int_{x_0}^{x_1} \alpha(x)U(x, t)\, dx \end{cases} \tag{9}$$

Introduce discrete time points $t^k = kh, k = 0, \dots, m$ where $mh = T$ denotes the final time for our simulation. Similarly, we introduce discrete ages $x_i = ih, i = 0, \dots, n$ where $nh \gg 1$ is large enough to cover the maximal life span of an individual. For the partial derivatives, we use the finite differences method where

$$\begin{cases} \dfrac{\partial}{\partial t} U(t^k, x_i) \cong \dfrac{U(t^{k+1}, x_i) - U(t^k, x_i)}{h} \\ \dfrac{\partial}{\partial x} U(t^k, x_i) \cong \dfrac{U(t^k, x_i) - U(t^k, x_{i-1})}{h} \end{cases} \tag{10}$$

Then we have the finite difference scheme of (6) as

$$U_i^{k+1} - U_{i-1}^k = -h\mu(x_i)U_i^k \tag{11}$$

where $U_i^k = U(x_i, t^k)$. The time step h can be chosen as one year. Replacing on the right-hand side U_i^k by U_{i-1}^k we get

$$U_i^{k+1} = U_{i-1}^k(1 - h\mu(x_i)) \tag{12}$$

where the factor $1 - h\mu(x_i)$ can be interpreted as the rate of survival of the age-cohort $i-1$ from the past year. In the sequel, we will use (12) instead of (11) since it allows for an easier interpretation.

The integral in the boundary condition can be discretized using any numerical quadrature, e.g.

$$U_0^k = h \sum_{i=i_0}^{i_1} \alpha(x_i)U_i^k \tag{13}$$

where $[i_0, i_1]$ denotes the age interval, when individuals reproduce. From (9), we obtain the following numerical scheme for all i as

$$\begin{cases} U_i^{k+1} = U_{i-1}^k(1 - h\mu(x_i)) \\ \qquad U_i^0 = U_0(x_i) \\ U_0^{k+1} = h \sum_{i=i_0}^{i_1} \alpha(x_i)U_i^{k+1} \end{cases} \tag{14}$$

We can apply this mathematical model to the population pyramid of a nation. The age structure model is used in the demography to project population structure based on age categories in the future. The population in a country is recorded periodically. In Indonesia, we do 10 years census (Sensus Populasi, SP) and 5 years census (Survey Penduduk Antar Sensus, SUPAS) [12]. Population data is categorized into five-yearly age groups namely 0-4 years, 5-9 years, 10-14 years, 15-19 years, 20-24 years, 25-29 years, 30-34 years, 35-39 years, 40-44 years, 45-49 years, 50-54 years, 55-59 years, 60-64 years, 65-69 years, 70-74 years, 75-79 years, 80-84 years, 85-89 years, 90-94 years, 95-99 years, and 100 years and above. This population data is displayed in the Population Pyramid or Age-Sex Pyramids. We refer to these graphs as pyramids because they are usually shaped like triangles, though as we will see shortly, population pyramids also take other shapes. Based on the Population Pyramid shapes, the policymakers can arrange strategies based on age-group in social, economic, educational, and human resource [13].

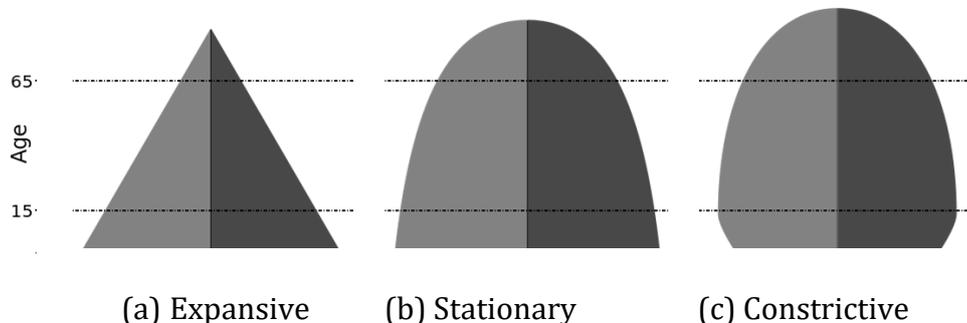(a) Expansive   (b) Stationary   (c) Constrictive

Figure 1. Population Pyramid Types

There are three types of population pyramids: expansive, constrictive, and stationary. Expansive population pyramids depict populations that have a larger

percentage of people in younger age groups. Populations with this shape usually have high fertility rates with lower life expectancies. Many third world countries have expansive population pyramids. Constrictive population pyramids are named so because they are constricted at the bottom. There is a lower percentage of younger people. Constrictive population pyramids show declining birth rates since each succeeding age group is getting smaller and smaller. The United States has a constrictive population pyramid. Stationary population pyramids are those that show a somewhat equal proportion of the population in each age group. There is not a decrease or increase in population; it is stable. Austria has a stationary population pyramid. These three types of population pyramid are illustrated in Figure 1.

The shape of a Population Pyramid can be used to interpret a population. For example, a pyramid with a very wide base and a narrow top section suggests a population with both high fertility and death rates. Whereas, a pyramid with a wider top half and a narrower base would suggest an aging population with low fertility rates. Population Pyramids can also be used to speculate a population's future development. An aging population that is not reproducing would eventually run into issues such as having enough offspring to care for the elderly. Other theories such as the "Youth Bulge" state that when there's a wide bulge around the 16-30 age range, particularly in males, this leads to social unrest, war, and terrorism.

Population Pyramids are ideal for detecting changes or differences in population patterns. Multiple Population Pyramids can be used to compare patterns across nations or selected population groups. This makes Population Pyramids useful for fields such as Ecology, Sociology, and Economics.

RESULTS AND DISCUSSION

The above model will be implemented in five countries which have different Population Pyramid shape, i.e. Indonesia, USA, Brazil, Russia, and Japan. Data that we used are taken from UN Data 2015. The death rate categories by the age-sex group can be shown in Figure 2, while Figure 3 shows the birth rate categories by the age-sex group.

Figure 2 shows the mortality rate in these five countries has increased with age. The mortality rate of males is higher compared to females. This indicated by a widening up and leaning to the left. The 90-94 age group in Brazil has the highest mortality rate compared to all age-group. The chance of death of age above 90 in Indonesia is zeros because of the data from the United Nation website for Indonesia only available up to age 85. Japan and the USA have a high level of elderly well-being which is indicated by the highest mortality rate at the age of over 100 years. The highest death rate in newborn age occurs in Indonesia while Japan gets the lowest death rate for zeros ages. This indicates that the knowledge level about new-born health related to sanitation and nutrition in Indonesia is lower than in Japan. Furthermore, newborns are very susceptible to any bacteria and viruses. Some people in Indonesia choose not to vaccinate their children for various reasons. This certainly will increase the chance that the baby died due to infectious diseases.
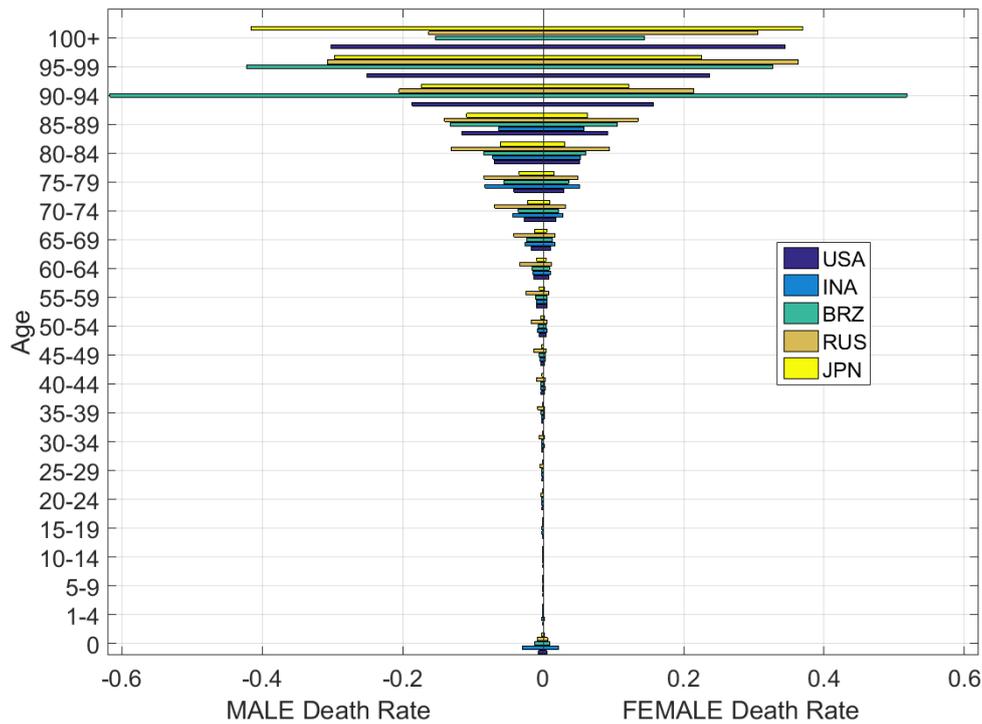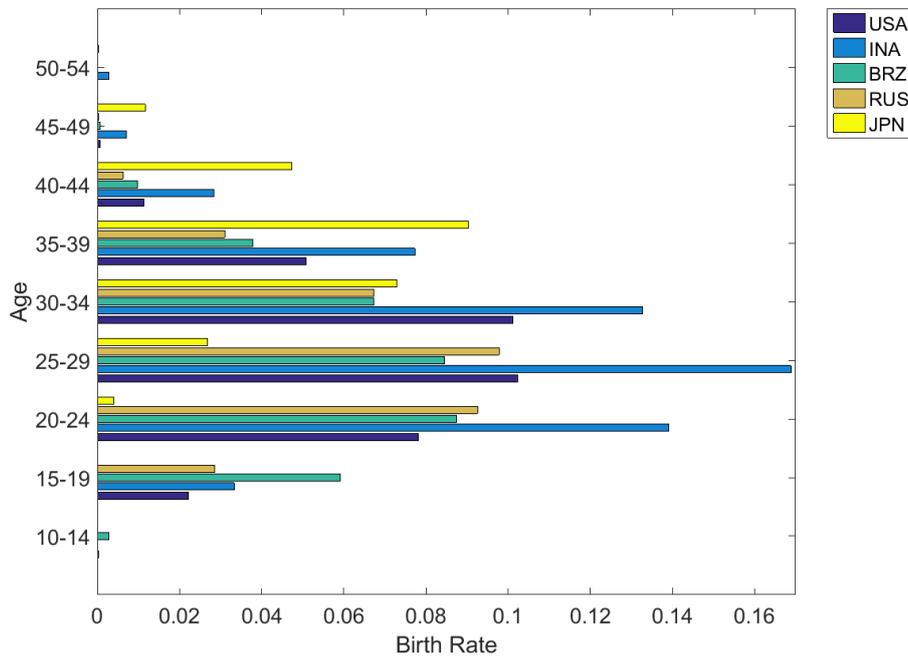
Figure 2. Death rate categories by age-sex group

Figure 3. Birth rate categories by mother's age group

Figure 3 shows the birth rate categories by mother's age group. Brazil has the youngest age of mothers at 10-14 years, while Indonesia has the oldest age of mothers, i.e. 50-54 years. If the USA and Russia have the youngest mother age at 15-19 years, Japan women begin to be a mother at 20-24 years old. The range of productive age in Japan

shifts 5 years older than women in Russia or the USA. This is because the majority of Japanese women prefer to prioritize career or education before 20 years and only decided to have children after 20 years. Moreover, this also becomes the reason why the highest birth rate in Japan given by 35-39 years age group. At this age-group, Japan women started to become economically established and start thinking about having a family. From the five countries, Indonesia has the highest birth rates in each age group. The highest maternal age groups for Indonesia, the USA, and Russia are in the same age group, i.e., 25-29 years.

(a) 2016 and 2021

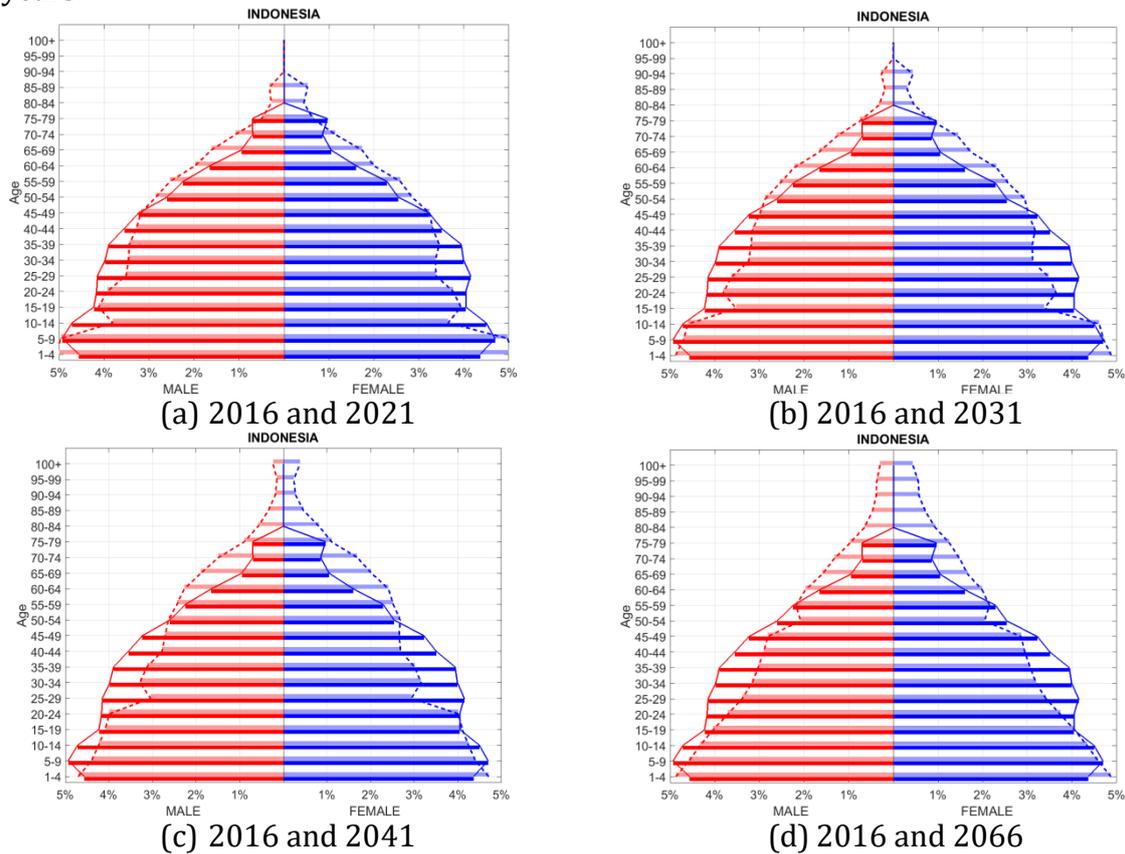(b) 2016 and 2031

(c) 2016 and 2041

(d) 2016 and 2066

Figure 4. Pyramid Population Indonesia (2016) and Prediction

Using the model (9) with the initial condition is the same with age structure data from the UN [14], we can make a pyramid population for each country. Beginning with the prediction population pyramid of Indonesia which shown in Figure 4. The population pyramid based on 2016 data is shown by a straight line while the prediction is shown by a dashed line. In 2021, the oldest age in Indonesia is 85-89 years. In the following ten years, the oldest age becomes 90-94 years. If group 5-9 year become the largest population in 2016, then in 2041 the largest population will be 20-24 years. From 2016 to 2041, the Pyramid has type as a stationary pyramid in which the percentage of adolescents and adults (25-69 years) almost the same. In 2066, the pyramid classified as an expansive population pyramid where the younger age groups have the largest percentage. In this year, Indonesia has high fertility rates with lower life expectancies.

Figure 5 shows the prediction of the population pyramid for Brazil. The population pyramid of Brazil in 2016 has a contracting type in which it has a lower percentage of the youngest age. Most age groups in 2016 are at the age of 30-34 years. This age group becomes the largest age group of the next 15 years where the individual age of this group becomes 45-49 years. In 2041, the largest age group is 55-59 years with a slimmer

pyramid shape. In 2066, the percentage of the old population is more than the young age. The shape of the pyramid that is formed is getting slimmer. If this continues, the population of Brazil will decrease from year to year.

(a) 2016 and 2026

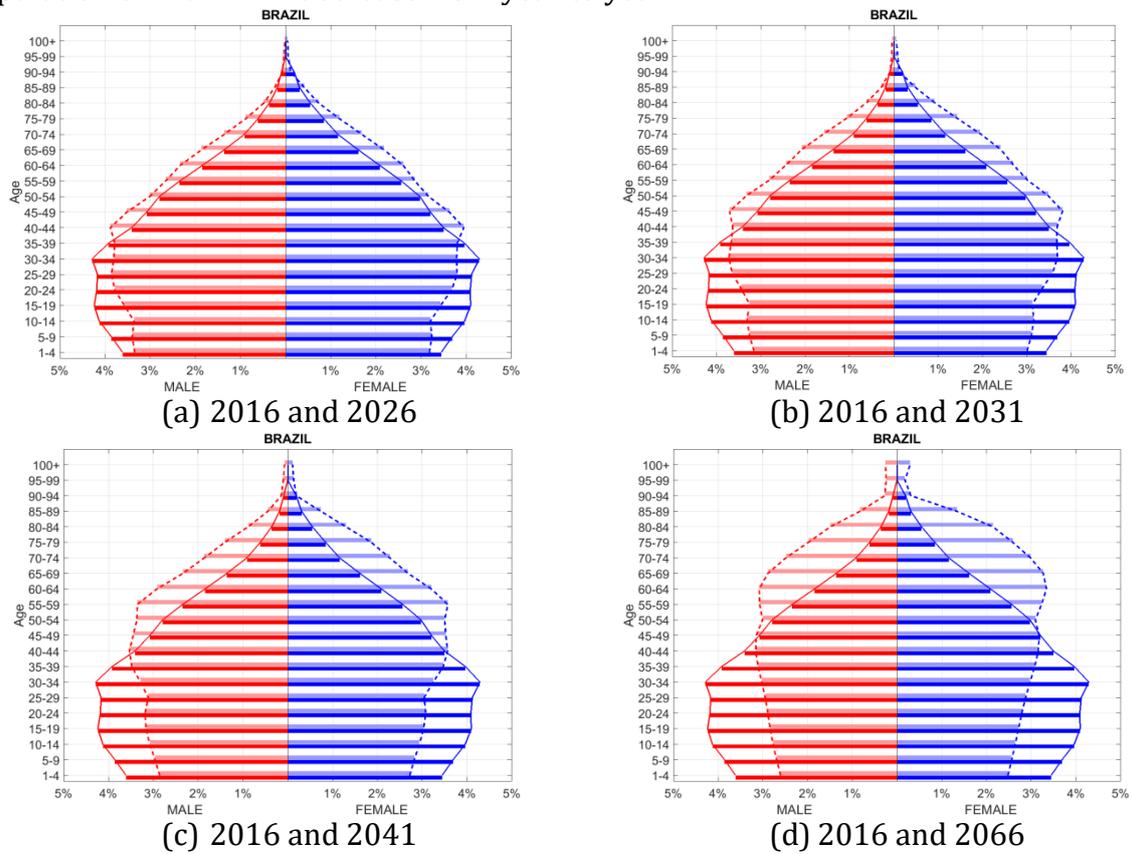(b) 2016 and 2031

(c) 2016 and 2041

(d) 2016 and 2066

Figure 5. Pyramid Population Brazil (2016) and Prediction

The other four countries before show a balanced composition of men and women in the population pyramid shape, but the Russian population pyramid shows an asymmetrical shape. The pyramid shape of the Russian population in 2016 is more inclined to the right which means that the percentage of the female population is higher than the male population. The population pyramid tends to contract type where the number of younger age is smaller. In 2066, the population pyramid shape is not as smooth as the shape of the pyramid of the other four countries which are shown in

Figure **8**.

(a) 2016 and 2026 — (b) 2016 and 2031
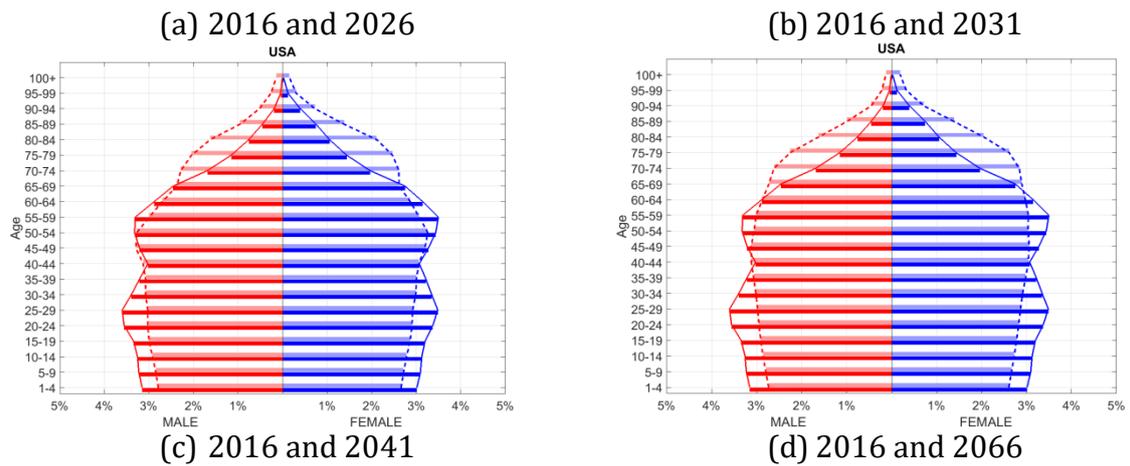
(c) 2016 and 2041 — (d) 2016 and 2066
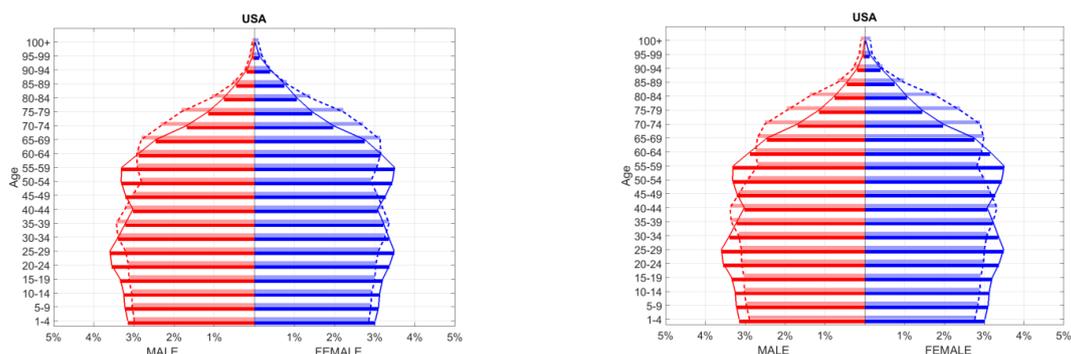
Figure 6 Pyramid Population USA (2016) and Prediction

Figure 6 shows the population pyramid prediction for the USA. From 2016 to 2066 the population pyramid has a stationary type where each age group has an equal proportion. The population is neither decrease nor increase. We called this population is stable. This characteristic is owned by developed countries where the government can control population growth well so that population growth still exists but does not explode.

The prediction of the Japanese population pyramid is shown in Figure 7. The pyramid type from 2016 to 2066 is a contracting type in which the number of young population is less than the old age. In 2016, the 40-44 and 65-69 age groups became the largest age group. This resulted in the year 2041, the largest age group was at the age of 65-69 years. The shape of the pyramid is increasingly symmetrical and smoother by 2066.
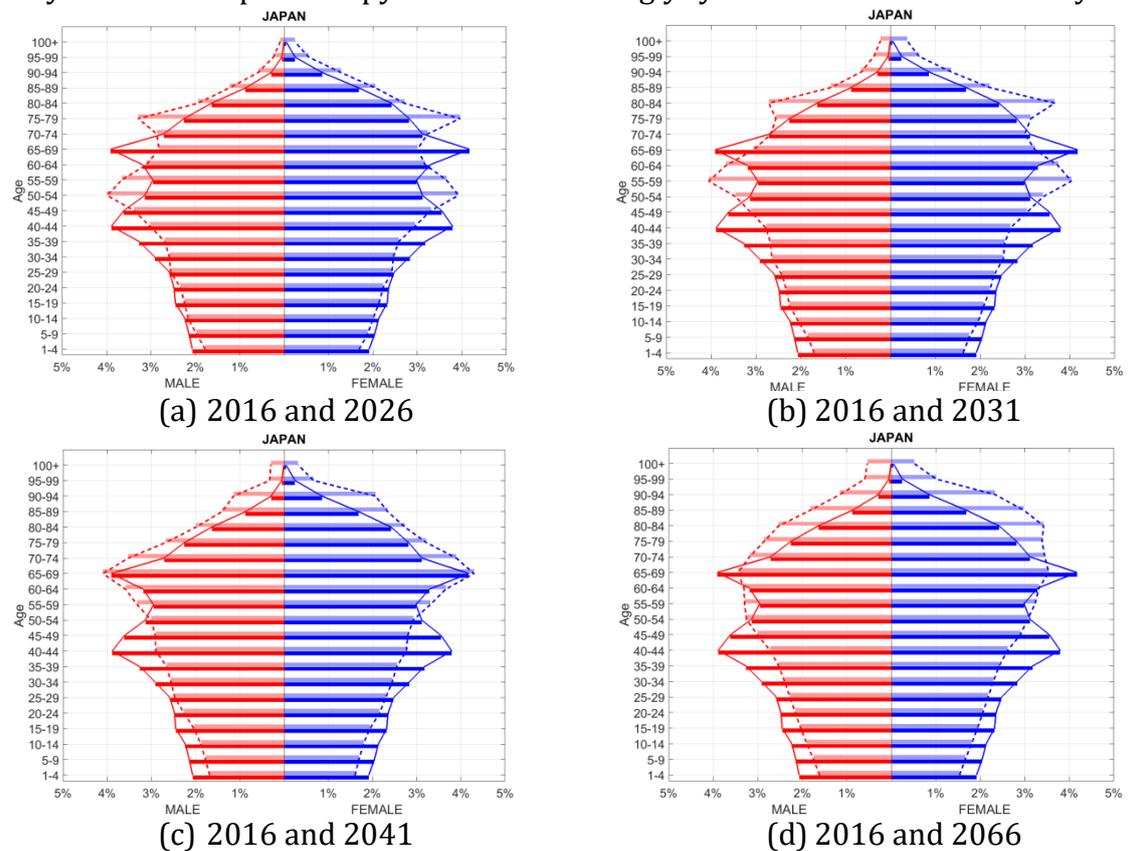

(a) 2016 and 2026 — (b) 2016 and 2031

(c) 2016 and 2041 — (d) 2016 and 2066

Figure 7. Pyramid Population Japan (2016) and Prediction

(a) 2016 and 2026

(b) 2016 and 2031
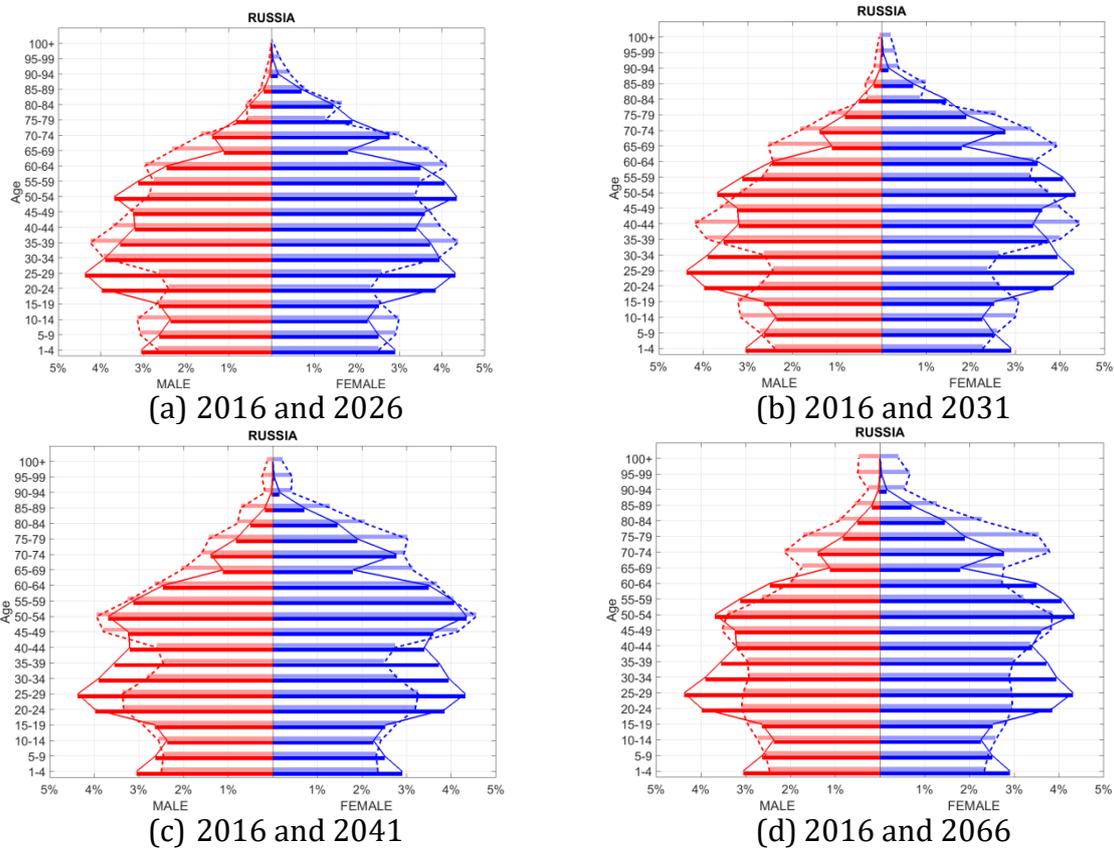
(c) 2016 and 2041

(d) 2016 and 2066

Figure 8. Pyramid Population Russia (2016) and Prediction

CONCLUSION

In this article, the age structure model for a population in a particular domain has been constructed. The proposed model has form as a linear one-order partial differential equation with non-negative initial conditions. The boundary condition is modeled from the cumulative birth rate of all reproductive age. The proposed model is implemented to predict the population pyramids from 5 countries i.e. Indonesia, Brazil, the USA, Japan, and Russia. These five countries were chosen because the population pyramid of the population represents the three types of population pyramid shape. From the simulation results, it can be seen that the age structure model can be used to predict the composition of a population in the future through the population pyramid. To improve the accuracy of the model, birth and death data in each age group must be complete and new for each year. This is because the rate of births and deaths in each year is very likely to change due to factors causing other than natural death. This model can also be developed by including the rate of immigration and emigration from a country. Data limitations are one challenge to predict population growth more accurately.

ACKNOWLEDGMENTS

REFERENCES

[1] Richmond, M. (2020, February 21). *Human Population Pyramids.* Retrieved February 21, 2020, from http://www/fsl.orst.edu/pnwerc/wrb/Atlas_web_compressed/.

[2] Christina, Kuttler. 2011. *Lecture Notes: Reaction-Diffusion equations with applications.* Munich: TU Munich.

[3] Auger, Pierre, Pierre Magal, and Shigui Ruan. 2008. *Structured population models in biology and epidemiology*. Vol. 1936. Berlin: Springer.

[4] Leal-Ramirez, Cecilia & Castillo, Oscar & Melin, Patricia & Rodríguez-Díaz, Antonio. 2011. *Simulation of the bird age-structured population growth based on an interval type-2 fuzzy cellular structure*. Inf. Sci.. 181. 519-535. 10.1016/j.ins.2010.10.011.

[5] Lande, Russell, and Steven Hecht Orzack. 1988. *Extinction dynamics of age-structured populations in a fluctuating environment.* Proceedings of the National Academy of Sciences 85.19 (1988): 7418-7421.

[6] Abia, L. M., O. Angulo, and J. C. López-Marcos. 2005. *Age-structured population models and their numerical solution*. Ecological Modelling 188.1 (2005): 112-136.

[7] Bacaër, Nicolas. 2011. *A short history of mathematical population dynamics*. Springer Science & Business Media.

[8] Malthus, Thomas Robert, Donald Winch, and Patricia James. 1992. *Malthus:'An Essay on the Principle of Population'*. Cambridge University Press.

[9] Verhulst, Pierre-François. *Deuxième mémoire sur la loi d'accroissement de la population*. 1847. Mémoires de l'académie royale des sciences, des lettres et des beaux-arts de Belgique 20 : 1-32.

[10] Lotka, Alfred J. 1931. *The structure of a growing population*. Human Biology 3.4: 459-493.

[11] Thomas, Gotz. 2018. *Compact Course: Numerics for Reaction-Diffusion Equations.* Depok: University of Indonesia.

[12] Hayes, Adrian, and Diahhadi Setyonaluri. 2015. *Taking advantage of the demographic dividend in Indonesia: A Brief Introduction to Theory and Practice*. United Nations Population Fund: New York, NY, USA.

[13] Pitoyo, A. J., et al. 2018. *System Dynamics Modeling of Indonesia Population Projection Model*. IOP Conference Series: Earth and Environmental Science. Vol. 145. No. 1. IOP Publishing.

[14] United Nation. 2016. *Demographic Yearbook 2016*. Retrieved from https://unstats.un.org/unsd/demographic-social/products/dyb/dyb_2016/