Probabilistic population projections for provincial levels in Indonesia

B Setiawan¹*, Sukamdi² and U Listyaningsih²

¹ Doctoral Student in Population Studies, Graduate School, Universitas Gadjah Mada, Jalan Teknika Utara, Pogung, Mlati, Sleman, Yogyakarta 55281
² Department of Environmental Geography, Faculty of Geography, Universitas Gadjah Mada, Sekip Utara, Bulaksumur, Yogyakarta 55281

*E-mail: budi.setiawan81@mail.ugm.ac.id

Abstract. This research aims to produce a probabilistic population projection for the provincial level in Indonesia. The base population uses the base population used by Bappenas, BPS, and UNFPA Indonesia in compiling the 2015-2045 Indonesian official population projection. One thousand trajectories of future E₀ values are projected using JPP with Asian countries as the prior distribution, while 1,000 trajectories of future TFR values are projected using BHM-based projections with Asian countries as the prior distribution. Meanwhile, the number and pattern of net internal migration are assumed to be constant during the projection period. The calculation process uses the bayesPop package in R software. The projection results show that the total population of the official projection is located at the predicted interval of the Bayesian population projection, except for the Kepulauan Riau.

1. Introductions

The National Development Planning Agency (BAPPENAS), Statistics Indonesia (BPS), and the United Nations Population Fund (UNFPA) Indonesia [1,2] still use a deterministic approach in making official population projections. The official population projection is a population projection prepared by BAPPENAS, BPS, and UNFPA Indonesia, which the Government of Indonesia officially uses for development planning purposes. There are two scenarios in making population projections, namely the policy scenario and the trend scenario. In the policy scenario, the total fertility rate (TFR) for the national level is assumed to be stagnant at 2.1 since 2020, while the infant mortality rate (IMR) for the national level is assumed to decrease by 3 percent a year from 2015 to 2030. For the provincial level, if the TFR has reached 2.1, then the TFR will be held at that value. In the trend scenario, the TFR and IMR at the national and provincial levels are projected to decline following the pattern of past TFR and IMR data using the logistical function. The scenario approach in preparing population projections has been criticized for lacking a probabilistic basis and leading to inconsistencies in quantifying future uncertainty [3–6].

In contrast to BAPPENAS, BPS, and UNFPA Indonesia, the United Nations (UN) has used a probabilistic approach in making population projections for all countries [7–10]. They implemented a Bayesian population projection initiated by Raftery et al. [11] and Raftery, Alkema, and Gerland [12]. In this model, TFR is projected probabilistically using the Bayesian hierarchical model (BHM) developed by Alkema et al. [13,14]. Meanwhile, life expectancy (E₀) is projected using a joint probabilistic projection (JPP) model for females' and males' life expectancy. This model was developed...
by Lalic and Raftery [15] and Raftery, Lalic, and Gerland [16] based on the BHM-based projection proposed by Chunn, Raftery, and Gerland [17] and Raftery et al. [18]. Meanwhile, migration in this model still uses a deterministic approach, which is based on the number of net international migration by age group and sex. The cohort component method is used to project population probabilistically based on these three components. The most recent UN projections have been published in World Population Prospects (WPP) 2019 [19].

The application of the Bayesian population projection above to create a probabilistic population projection at the subnational level is scarce. So far, Adil [20] initiated the implementation of this projection model to project population at the subnational level in Pakistan. In line with Adil [20], this research aims to make a probabilistic population projection at the provincial level in Indonesia using a Bayesian population projection.

2. Materials and Methods

2.1. Base population.

The base population uses the base population used by Bappenas, BPS, and UNFPA Indonesia in compiling the 2015-2045 Indonesian official population projection [1]. This base population is the population according to the five-year age group and sex from the 2015 Inter-Census Population Survey (IPS 2015), which has been tidied up and shifted to the middle of the year using the BPSTRNG.xls and MOVEPOP.xls templates. Both of these templates are included in the Rural-Urban Projection (RUP) software [2]. Figure 1 shows the population from IPS 2015 and the base population for Jawa Barat, the most populous province in Indonesia.

![Figure 1. Population from IPS 2015 and base population for Jawa Barat.](image)

2.2. Probabilistic projections for \( E_0 \).

The trajectory of the future \( E_0 \) values uses the 1,000 trajectories of the future \( E_0 \) value generated by Setiawan, Sukamdi, and Listyaningsih [21]. They produced 1,000 future \( E_0 \) value trajectories for females and males using the JPP with Asian countries as the prior distribution. Figure 2 shows the 1,000 future \( E_0 \) value trajectories for SR Yogyakarta, the province with the highest life expectancy in Indonesia.
2.3. Probabilistic projections for TFR.

The trajectory of the future TFR values uses the 1,000 trajectories of the future TFR values generated by Setiawan, Sukamdi, and Listyaningsih [22]. They produce 1,000 trajectories of future TFR values using BHM-based projections with Asian countries as the prior distribution. Figure 3 shows the 1,000 trajectories of future TFR values for Nusa Tenggara Timur, the province with the highest TFR in Indonesia.
2.4. **Internal migration.**
Internal migration uses net internal migration from IPS 2015. Net internal migration is grouped according to five-year age group and sex. The number and pattern of net internal migration are assumed to be constant over the projection period.

2.5. **Bayesian population projections.**
This research adopts the Bayesian population projection developed by Raftery et al. [11] and Raftery, Alkema, et al. [12]. Bayesian population projection is a probabilistic population projection using the cohort component method to generate a number of future population trajectories based on a number of trajectories of future $Eo$ and TFR values. In this research, the Bayesian population projection produces 1,000 future population trajectories based on 1,000 trajectories of future $Eo$ and TFR values. The number of those trajectories refers to the number of trajectories used by Hana Ševčíková to project $Eo$ and TFR probabilistically for all countries on the website BayesPop, Probabilistic Population Projections (https://bayespop.csss.washington.edu/) and considering laptop specification which researchers use.

The future $Eo$ value trajectories are projected using JPP with Asian countries as the prior distribution, while the future TFR value trajectories are projected using BHM-based projection with Asian countries as the prior distribution. Age-specific mortality rate ($m_8$) was projected using the modified Lee-Carter method based on the trajectory of future $Eo$ values for female and male and coherent Lee-Carter parameter estimates [23]. The coherent Lee-Carter parameter estimate is based on Indonesia's past $m_8$ values obtained from the wpp2019 package in R software [24]. Future trajectories of TFR values were converted to age-specific fertility rate (ASFR) using Indonesia's proportionate age-specific fertility rate (PASFR) projection obtained from the wpp2019 package in R software [24]. Meanwhile, migration in this research still uses a deterministic approach, which is based on the number of net internal migrants according to the five-year age group and sex. The calculation process uses the bayesPop package v.8.1-3 in R software [25].

3. **Results and Discussions**

3.1. **Bayesian population projections at the provincial level.**
Table 1 shows the total population from the official projections and the Bayesian population projections for the provincial level in Indonesia using the cohort component method. At the beginning of the projection period, most provinces show the total population of the official projection to be higher than the median value of the Bayesian population projection. These provinces are Riau, Jambi, Sumatera Selatan, SCR Jakarta, SR Yogyakarta, Jawa Timur, Banten, Bali, Nusa Tenggara Barat, Kalimantan Tengah, Kalimantan Selatan, Sulawesi Utara, Sulawesi Tengah, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, and Papua Barat. In contrast, other provinces show the total population of the official projection to be lower than the median value of the Bayesian population projection. However, the total population of these provinces still lies within the 80 percent prediction interval of the Bayesian population projection. The only province where the total population of the official projection lies outside the 95 percent prediction interval of the Bayesian population projection is Kepulauan Riau.

At the end of the projection period, almost all provinces show that the total population of the official projection is higher than the median value of the Bayesian population projection, except for Jawa Tengah, Nusa Tenggara Timur, and Papua. These three provinces show the total population of the official projection to be lower than the median value of the Bayesian population projection even though they are still located within the 80 percent prediction interval of the Bayesian population projection. Meanwhile, Kepulauan Riau is still the only province where the total population of the official projection lies outside the 95 percent prediction interval of the Bayesian population projection.
| Provinces               | 2020 Median | Lower95 | Lower80 | Upper80 | Upper95 Median | Lower95 | Lower80 | Upper80 | Upper95 |
|-------------------------|-------------|---------|---------|---------|----------------|---------|---------|---------|---------|
| Aceh                    | 5,388       | 5,402   | 5,321   | 5,347   | 5,453          | 5,482   | 6,906   | 6,824   | 6,095   | 6,326 |
| Sumatera Utara          | 14,798      | 14,579  | 14,560  | 14,660  | 14,949         | 15,025  | 17,656  | 17,213  | 15,292  | 15,970 |
| Sumatera Barat          | 5,546       | 5,560   | 5,475   | 5,498   | 5,616          | 5,640   | 6,900   | 6,850   | 6,155   | 6,370 |
| Riau                    | 6,951       | 6,918   | 6,815   | 6,853   | 6,888          | 7,023   | 9,727   | 9,247   | 8,344   | 8,635 |
| Jambi                   | 3,604       | 3,597   | 3,545   | 3,562   | 3,635          | 3,653   | 4,250   | 4,124   | 3,678   | 3,827 |
| Sumatera Selatan        | 8,601       | 8,557   | 8,430   | 8,473   | 8,648          | 8,689   | 10,597  | 10,088  | 9,049   | 9,374 |
| Bengkulu                | 1,994       | 1,996   | 1,968   | 1,977   | 2,015          | 2,026   | 2,392   | 2,261   | 2,115   | 2,208 |
| Lampung                 | 8,535       | 8,536   | 8,407   | 8,451   | 8,615          | 8,655   | 9,727   | 9,602   | 8,643   | 8,986 |
| Bangka Belitung         | 1,470       | 1,470   | 1,449   | 1,456   | 1,484          | 1,493   | 1,844   | 1,811   | 1,628   | 1,689 |
| Kepulauan Riau          | 2,310       | 2,269   | 2,232   | 2,245   | 2,292          | 2,303   | 4,342   | 3,635   | 3,280   | 3,413 |
| SCR Jakarta             | 10,576      | 10,507  | 10,399  | 10,408  | 10,664         | 10,662  | 11,239  | 10,784  | 9,488   | 9,936 |
| Jawa Barat              | 49,565      | 49,830  | 49,046  | 49,432  | 50,334         | 50,609  | 60,315  | 59,909  | 54,007  | 55,866 |
| Jawa Tengah             | 34,738      | 34,941  | 34,414  | 34,594  | 35,257         | 35,428  | 36,874  | 36,893  | 32,993  | 34,248 |
| SR Yogyakarta           | 3,919       | 3,892   | 3,837   | 3,855   | 3,929          | 3,947   | 5,217   | 4,873   | 4,357   | 4,532 |
| Jawa Timur              | 39,956      | 39,743  | 39,148  | 39,359  | 40,103         | 40,317  | 41,615  | 40,037  | 35,885  | 37,181 |
| Banten                  | 12,895      | 12,859  | 12,651  | 12,720  | 12,984         | 13,090  | 16,570  | 15,863  | 14,214  | 14,789 |
| Bali                    | 4,414       | 4,411   | 4,344   | 4,371   | 4,455          | 4,480   | 5,535   | 5,365   | 4,761   | 5,018 |
| Nusa Tenggara Barat     | 5,226       | 5,207   | 5,126   | 5,152   | 5,264          | 5,301   | 6,710   | 6,517   | 5,840   | 6,066 |
| Nusa Tenggara Timur     | 5,513       | 5,536   | 5,443   | 5,473   | 5,599          | 5,634   | 7,071   | 7,194   | 6,291   | 6,622 |
| Kalimantan Barat        | 5,105       | 5,125   | 5,052   | 5,074   | 5,177          | 5,205   | 6,196   | 6,179   | 5,495   | 5,731 |
| Kalimantan Tengah       | 2,686       | 2,679   | 2,637   | 2,651   | 2,704          | 2,717   | 3,387   | 3,221   | 2,893   | 3,001 |
| Kalimantan Selatan      | 4,269       | 4,249   | 4,189   | 4,208   | 4,293          | 4,317   | 5,327   | 5,166   | 4,600   | 4,769 |
| Kalimantan Timur        | 4,373       | 4,388   | 4,328   | 4,349   | 4,432          | 4,452   | 5,517   | 5,482   | 4,930   | 5,113 |
| Sulawesi Utara          | 2,513       | 2,508   | 2,472   | 2,484   | 2,533          | 2,546   | 2,772   | 2,696   | 2,424   | 2,518 |
| Sulawesi Tengah         | 3,082       | 3,055   | 3,010   | 3,024   | 3,087          | 3,101   | 3,940   | 3,643   | 3,267   | 3,368 |
| Sulawesi Selatan        | 8,889       | 8,927   | 8,784   | 8,831   | 9,010          | 9,073   | 10,004  | 9,856   | 8,714   | 9,130 |
| Sulawesi Tenggara       | 2,704       | 2,689   | 2,643   | 2,659   | 2,716          | 2,730   | 3,583   | 3,363   | 2,984   | 3,124 |
| Gorontalo               | 1,186       | 1,182   | 1,165   | 1,170   | 1,184          | 1,200   | 1,349   | 1,266   | 1,130   | 1,179 |
| Sulawesi Barat          | 1,378       | 1,376   | 1,356   | 1,362   | 1,390          | 1,398   | 1,758   | 1,730   | 1,545   | 1,604 |
| Maluku                  | 1,787       | 1,781   | 1,754   | 1,763   | 1,799          | 1,810   | 2,155   | 2,034   | 1,813   | 1,889 |
| Maluku Utara            | 1,252       | 1,255   | 1,236   | 1,242   | 1,268          | 1,275   | 1,615   | 1,588   | 1,420   | 1,471 |
| Papua Barat             | 986         | 986     | 971     | 976     | 996            | 1,001   | 1,603   | 1,479   | 1,338   | 1,388 |

Source: BAPPENAS, BPS, & UNFPA Indonesia [1]
3.2. **Aggregation at the national level.**

In the bayesPop package, there are two aggregation methods, namely *country-based* and *region-based* methods [25]. Aggregation in the country-based method is carried out by simply adding up the number of populations in each trajectory in all countries in the region concerned. The aggregation in the region-based method is created using the cohort component method, similar to pop.predict, but its function operates on the aggregate input component. In practice, Ševčíková and Raftery [25] found that, when projecting a pool of countries whose demographic histories do not coincide, region-based methods tend to overestimate uncertainty, often providing too wide a prediction interval. Therefore, the aggregation in this research uses country-based methods.

There are two scenarios in the official projection, namely the policy scenario and the trend scenario. In 2020, the total population of the policy-based projection was slightly lower than the median value of the Bayesian population projection (Table 2). After that, the total population from the policy-based projection exceeds the median value of the Bayesian population projection. The total population difference between these two projections will reach 7.7 million at the end of the projection period. Even so, the total population from the policy-based projection still lies within the 80 percent prediction interval of the Bayesian population projection. Meanwhile, the total population from the trend-based projection is similar to the median value of the Bayesian population projection. The difference in total population between the two projections ranges from 5 to 406 thousand.

| Table 2. Bayesian population projections of Indonesia, 2015-2045 (thousands). |
|---------------------------------------------------------------|
| **Source of Projections** | **2015** | **2020** | **2025** | **2030** | **2035** | **2040** | **2045** |
| Bayesian population projections | | | | | | | |
| Median | 255,588 | 269,629 | 282,047 | 292,566 | 300,990 | 307,226 | 311,246 |
| Lower 80 | 255,588 | 268,533 | 279,756 | 289,048 | 296,041 | 300,365 | 302,380 |
| Upper 80 | 255,588 | 270,740 | 284,336 | 296,348 | 306,409 | 314,477 | 320,505 |
| Lower 95 | 255,588 | 268,011 | 278,564 | 287,368 | 293,701 | 297,335 | 298,605 |
| Upper 95 | 255,588 | 271,284 | 285,729 | 298,028 | 309,067 | 318,175 | 325,638 |
| Official projection with two scenarios \( ^a \) | | | | | | | |
| Policy Scenario | 255,588 | 269,603 | 282,455 | 294,116 | 304,212 | 312,506 | 318,961 |
| Trend Scenario | 255,588 | 269,634 | 282,021 | 292,540 | 301,017 | 307,386 | 311,652 |
| UN's projections \( ^b \) | | | | | | | |
| Median | 273,524 | 287,090 | 299,198 | 309,765 | 318,638 | 325,705 | 330,905 |
| Source: |
| \( ^a \) BAPPENAS, BPS, & UNFPA Indonesia [1] |
| \( ^b \) United Nations Department of Economic and Social Affairs Population Division [19] |

There is a significant difference in total population between the median values of the two Bayesian population projections, namely the UN projection and this research. During the projection period, the total population of the median value of the UN projection lies outside the 95 percent prediction interval of the Bayesian population projection in this research (Figure 4). This difference is mainly due to the difference between the total population, which is used as the base population. The total population used as the base population of the UN projection is 273.5 million. Meanwhile, the total population used as the base population of this research is 255.6 million. The difference between the total population of these two projections is relatively constant over the projection period.
4. Conclusions
This research shows that Bayesian population projections can be relied on to make probabilistic population projections at the provincial level in Indonesia. The uncertainty in the future is quantified by means of a prediction interval where the result of the official projection lies within the prediction interval. Therefore, the official population projection based on the results of the 2020 Population Census is recommended to apply the Bayesian population projection.

However, there are several limitations to this research. First, to convert future TFR projections into ASFR, still use Indonesia's future PASFR. Second, the bayesPop package only provides the modified Lee-Carter method to project the population at the subnational level. Therefore, to convert future E0 projections into m0, still use Indonesia's past m0. Third, migration is the only projection component that still uses a deterministic approach. Therefore, in future studies, this component will be projected using a probabilistic approach.

Acknowledgments
We thank Prof. Adrian Raftery and Dr. Hana Ševčíková (University of Washington), pioneers of the Bayesian population projection for all countries, for sharing their knowledge on this subject at the 2019 Applied Bayesian Statistics School in Como, Italy. In particular, Dr. Hana Ševčíková for overcoming technical problems that arise in the calculation process using the bayesPop package. We also thank Dr. Patrick Gerland (Population Estimates and Projections Section, Population Trends and Analysis Branch, Population Division) for explaining the estimation techniques used by the UN Population Division in estimating the birth and death rates for the 1950-2020 period and explaining the demographic balance equation in the Bayesian population projection of WPP 2019 for Indonesia.

References
[1] Kementerian PPN/Bappenas, Badan Pusat Statistik and United Nations Population Fund Indonesia 2018 *Proyeksi Penduduk Indonesia 2015-2045: Hasil SUPAS 2015* (Jakarta: Badan Pusat Statistik)
[2] Kementerian PPN/Bappenas, Badan Pusat Statistik and United Nations Population Fund Indonesia 2018 *Technical Note: Proyeksi Penduduk Indonesia 2015-2045: Hasil SUPAS 2015* (Jakarta: Badan Pusat Statistik)
[3] Lee R D and Tuljapurkar S 1994 Stochastic Population Forecasts for the United States: Beyond
High, Medium, and Low J. Am. Stat. Assoc. 89 1175–89

[4] Lee R D 1999 Probabilistic Approaches to Population Forecasting Frontiers of Population Forecasting, Supplement to Population and Development Review, vol. 24 ed W Lutz, J W Vaupel and D A Ahlburg (New York: Population Council) pp 156–90

[5] National Research Council 2000 Beyond Six Billion: Forecasting the World's Population ed J Bongaarts and R A Bulatao (Washington, D.C.: National Academy Press)

[6] Lutz W and Goldstein J R 2004 Introduction: How to Deal with Uncertainty in Population Forecasting? Int. Stat. Rev. 72 1–4

[7] United Nations Department of Economic and Social Affairs Population Division 2014 World Population Prospects: The 2012 Revision, Methodology of the United Nations Population Estimates and Projections

[8] United Nations Department of Economic and Social Affairs Population Division 2015 World Population Prospects: The 2015 Revision, Methodology of the United Nations Population Estimates and Projections

[9] United Nations Department of Economic and Social Affairs Population Division 2017 World Population Prospects: The 2017 Revision, Methodology of the United Nations Population Estimates and Projections (New York)

[10] United Nations Department of Economic and Social Affairs Population Division 2019 World Population Prospects 2019: Methodology of the United Nations population estimates and projections

[11] Raftery A E, Li N, Ševčíková H, Gerland P and Heilig G K 2012 Bayesian Probabilistic Population Projections for All Countries Proc. Natl. Acad. Sci. 109 13915–21

[12] Raftery A E, Alkema L and Gerland P 2014 Bayesian Population Projections for the United Nations Stat. Sci. 29 58–68

[13] Alkema L, Raftery A E, Gerland P, Clark S J, Pelletier F and Buettner T 2010 Probabilistic Projections of the Total Fertility Rate for All Countries (Washington, D.C.)

[14] Alkema L, Raftery A E, Gerland P, Clark S J, Pelletier F, Buettner T and Heilig G K 2011 Probabilistic Projections of the Total Fertility Rate for All Countries Demography 48 815–39

[15] Lalic N and Raftery A E 2012 Joint probabilistic projection of female and male life expectancy Demogr. Res. 30 795–822

[16] Raftery A E, Lalic N and Gerland P 2014 Joint Probabilistic Projection of Female and Male Life Expectancy Demogr. Res. 30 795–822

[17] Chunn J L, Raftery A E and Gerland P 2010 Bayesian Probabilistic Projections of Life Expectancy for All Countries (Washington, D.C.)

[18] Raftery A E, Chunn J L, Gerland P and Ševčíková H 2013 Bayesian Probabilistic Projections of Life Expectancy for All Countries Demography 50 777–801

[19] United Nations Department of Economic and Social Affairs Population Division 2019 World Population Prospects 2019, Volume I: Comprehensive Tables

[20] Adil M 2017 Projecting Pakistan Population with a Bayesian Hierarchical approach (Università degli Studi di Padova)

[21] Setiawan B, Sukamdi and Listyaningsih U 2020 Probabilistic Projections of Life Expectancy for Indonesia and Its Regions using the Joint Probabilistic Projection Model

[22] Setiawan B, Sukamdi and Listyaningsih U 2020 Probabilistic Projections of Total Fertility Rates for Indonesia and Its Regions using the Bayesian Hierarchical Model

[23] Ševčíková H, Kantorová V, Gerland P and Raftery A E 2016 Age-Specific Mortality and Fertility Rates for Probabilistic Population Projections Dynamic Demographic Analysis ed R Schoen (Springer, Cham) pp 285–310

[24] United Nations Population Division 2019 Package "wpp2019"

[25] Ševčíková H and Raftery A E 2016 bayesPop: Probabilistic Population Projections J. Stat. Softw. 75 1–29