

Probabilistic Projection of Subnational Life Expectancy

Hana Ševčíková, Adrian E. Raftery
University of Washington

1 Introduction

In 2015 for the first time, the United Nations Population Division issued official probabilistic population projections for all countries (United Nations 2015), using the methodology described by Raftery et al. (2012). The two key components that are forecasted probabilistically are the total fertility rate (TFR) and life expectancy at birth (e_0). They are both based on Bayesian hierarchical models (BHM) that are tailored to work well at the national level.

At the subnational level, e.g. provinces, states, counties, regions (which we will refer to as regions for the purposes of this article), probabilistic population projections are of great interest to national and local governments for planning, policy and decision-making (Rayer et al. 2009). In Ševčíková et al. (2018) a generic method was developed for projecting subnational TFR probabilistically, based on the BHM for national TFR.

Here we introduce three methods for forecasting regional mortality that are probabilistic, easy to implement, and are based on the established BHM for national life expectancy. They yield probabilistic forecasts of life expectancy at birth, each of which can be converted to age-specific mortality patterns, yielding a probability distribution of mortality rates. Such age-specific mortality forecasts can be further transformed, providing full probability distributions of any future life table quantity. Reporting mortality-related measures while taking account of uncertainty can become critical, especially in times of rapid population aging, for example for policy makers who deal with designing sustainable pension systems.

2 Data

We use data on subnational female life expectancy at birth, e_0 , for 447 regions from 29 countries spanning five continents. They were collected from different sources, including EuroStat, Statistics Canada, the Australian Bureau of Statistics, and the Institute for Health Metrics and Evaluation (IHME). We put the data on a similar time scale to the national projections, resulting in a dataset for 5-year time periods, with the earliest being 1920-1925 and the latest being 2010-2015. The majority of the regions are from European countries. However, the European time series are on average shorter than the time series from other continents. For national data, we use the time series from World Population Prospects (WPP) 2017 (United Nations 2017).

3 Methods

3.1 Methods Overview

For generating projections of subnational e_0 , we seek a method that is probabilistic, is based on the established national methodology used by the UN (Raftery et al. 2013), works well for most countries, is simple, and yields correlations between regions that are similar to the correlations in the observed data.

We introduce three methods that satisfy at least most of these requirements, and we compare them to two other baseline methods. Two of the new methods are adaptations of simple deterministic methods to a probabilistic framework. The third method is an extension to take into account correlation between regions.

The starting point for all three of the new methods is projections at the national level. First, a Bayesian hierarchical model is used to produce a set of trajectories of future female e_0 from its posterior predictive distribution (Raftery et al. 2013). Then a female-male gap model is used to generate trajectories of male e_0 (Raftery et al. 2014). To project at the subnational level, we apply the methods described below to each national trajectory, yielding probabilistic projections for each region.

In the following, $e_{c,t}^{(C)}$ denotes the life expectancy of country c at time t , and $e_{r,c,t}^{(R)}$ denotes the life expectancy of region r of country c at time t . In all three methods we model $e_{r,c,t}^{(R)}$ as

$$e_{r,c,t}^{(R)} = f(e_{c,t}^{(C)}, \alpha_{r,c}), \quad (1)$$

namely that the regional life expectancy is a function of the national life expectancy and a regional term $\alpha_{r,c}$. The difference between the three methods lies in the definition of the function $f(\cdot, \cdot)$ and how the regional term $\alpha_{r,c}$ is modeled.

Scale method: The Scale method follows Wilson (2018) and defines $\alpha_{r,c}$ as a time-invariant factor, i.e.

$$e_{r,c,t}^{(R)} = \alpha_{r,c} e_{c,t}^{(C)}, \quad (2)$$

where $\alpha_{r,c} = e_{r,c,t_0}^{(R)} / e_{c,t_0}^{(C)}$ is derived from the last observed time period t_0 .

The method is implemented by taking the i -th trajectory simulated from the posterior predictive distribution of future national life expectancy and multiplying it by the scale factor $\alpha_{r,c}$. This is done for each trajectory, yielding a set of trajectories of future regional life expectancy, representing its posterior predictive distribution. Thus if $e_{c,t,i}^{(C)}$ is the simulated value of future national life expectancy from the i trajectory, then the corresponding simulated value of regional life expectancy is $e_{r,c,t,i}^{(R)} = \alpha_{r,c} e_{c,t,i}^{(C)}$. Note that in this method, the scale factor $\alpha_{r,c}$ is constant across trajectories.

This method is simple and probabilistic. However, it has the drawback of yielding perfectly parallel trajectories with no crossovers between regions, whereas in fact such crossovers do occur.

Shift method: Here, the regional term is defined as a time invariant additive constant, i.e.

$$e_{r,c,t}^{(R)} = e_{c,t}^{(C)} + \alpha_{r,c}, \quad (3)$$

where $\alpha_{r,c} = e_{r,c,t_0}^{(R)} - e_{c,t_0}^{(C)}$ with t_0 being the last observed time period. It is implemented similarly to the Scale method. Like the Scale method, this method is simple and probabilistic, but yields parallel trajectories, not allowing for the possibility of crossovers.

AR(1) method: This method tries to rectify the drawback of the above methods that they produce perfectly parallel trajectories and thus do not account for the possibility of crossovers between regions. To do this, we leverage a perhaps surprising conclusion from Ševčíková et al. (2018) that subnational TFR can be well modeled using a simple scale method with a scale factor that is stochastic and follows a first-order autoregressive, or AR(1) model.

We propose a method for projecting subnational e_0 in a similar way, extending the Shift method. We take the regional term $\alpha_{r,c,t}$ to be time dependent and additive, i.e.

$$e_{r,c,t}^{(R)} = e_{c,t}^{(C)} + \alpha_{r,c,t}. \quad (4)$$

It is modeled using a first-order autoregressive, or AR(1), process:

$$\alpha_{r,c,t} = \rho \alpha_{r,c,t-1} + \varepsilon_{r,c,t}, \quad \text{with } \varepsilon_{r,c,t} \stackrel{\text{ind}}{\sim} N(0, \sigma_{c,t}^2), \quad (5)$$

where ρ is the autoregressive parameter, which is constant across all countries and regions and constrained to lie between 0 and 1, and $\sigma_{c,t}^2$ is the residual variance, which is allowed to vary between countries.

This is implemented similarly to the Shift method, with the addition of simulating the regional term, which is no longer taken to be constant in time but instead is allowed to vary randomly across time and between trajectories. Specifically, we simulate the i -th trajectory of the future values of the stochastic regional term, denoted by $\alpha_{r,c,t,i}$, as follows:

$$\begin{aligned} \alpha_{r,c,t_0,i} &= e_{r,c,t_0}^{(R)} - e_{c,t_0}^{(C)} \\ \alpha_{r,c,t,i} &= \rho \alpha_{r,c,t-1,i} + \varepsilon_{r,c,t,i} \text{ for } t > t_0, \end{aligned} \quad (6)$$

where the $\varepsilon_{r,c,t,i}$ are simulated as independently distributed random variables from a $N(0, \sigma_{c,t}^2)$ distribution. Then the i -th trajectory from the posterior predictive distribution of regional life expectancy is $e_{r,c,t,i}^{(R)} = e_{c,t,i}^{(C)} + \alpha_{r,c,t,i}$.

We set $\sigma_{c,t}^2$ to be a function of $e_{c,t}^{(C)}$, namely:

$$\sigma_{c,t}^2 = \begin{cases} a + b(e_{c,t}^{(C)} - U) & e_{c,t}^{(C)} < U \\ a & e_{c,t}^{(C)} \geq U, \end{cases} \quad (7)$$

where U is the value of life expectancy at which the variance switches from being a linearly increasing function of the country-specific life expectancy $e_{c,t}^{(C)}$ with slope b , to a constant, a . The parameters a , b and U are taken as constant across countries and regions.

Using the observed data, we estimate $\hat{\rho} = 0.95$, $U = 82.5$, $a = 0.0482$, and $b = -0.0154$.

4 Results

To predict female e_0 at the national level, we simulated a BHM, using the bayesLife R package (Ševčíková et al. 2019), resulting in 1,000 trajectories of life expectancy for each country, from 2020 to 2100. Then we applied the three methods above to each trajectory to obtain 1,000 future trajectories for each region in our data.

4.1 Predictive Distribution of Life Expectancy

We now present the marginal predictive distribution from the AR(1) and the Shift methods for three states in the USA, see Figure 1. Mississippi (on the left), currently has the lowest female life expectancy in the country. The distribution from the AR(1) method (red shaded area) is predicted to stay below the national distribution (grey area), with the gap between the medians slowly decreasing. The Shift method (yellow shaded area) keeps the gap between the medians constant, and thus the projection is lower than for AR(1). Virginia (in the middle), has female e_0 currently similar to the national one, and the projections from both methods follow the national distribution with slightly higher uncertainty. Finally Hawaii, the state with the highest female e_0 , is projected to stay mostly above the national distribution with a slow decrease of the gap between the two distributions for the AR(1), and a constant gap for the Shift method. These three states illustrate the types of results obtained for all regions in our data. Note that the Scale method yields very similar marginal distributions to the Shift method. Thus for clarity, we omit it in Figure 1.

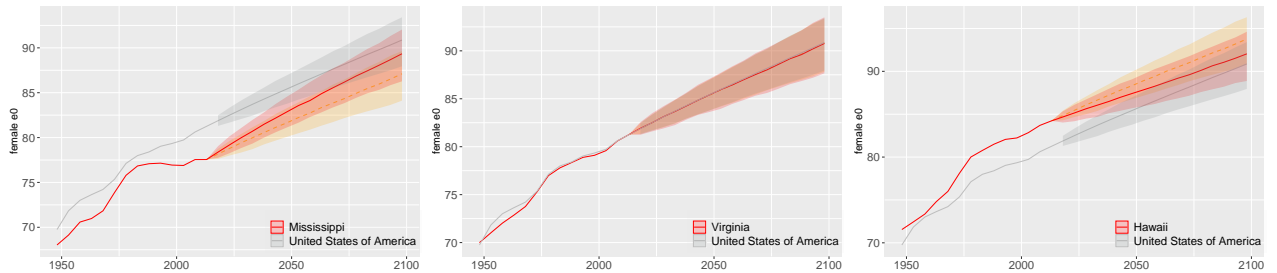


Figure 1: Predictive distribution of female e_0 for Mississippi, Virginia and Hawaii. Observed data and projected medians are shown as solid lines, while the 80% probability interval is shown as a shaded area. The subnational results from the AR(1) method are shown in red, while the national data and projections are shown in grey. Results from the Shift method are shown in light yellow with its median as dashed line.

4.2 Validation

This section compares results from the three methods to two additional methods:

National: The regional projections are set to the national projections, so that $e_{r,c,t,i}^{(R)} = e_{c,t,i}^{(C)}$ for all trajectories i .

Persistence: The regional projections stay the same over time, i.e. $e_{r,c,t}^{(R)} = e_{r,c,t_0}^{(R)}$, yielding deterministic projections.

To assess the marginal predictions of e_0 made by our methods, we performed an out-of-sample validation exercise for the predicted $e^{(R)}$. We withheld data for two to five time periods, corresponding to 10 to 25 years. We did this for both the national model and the subnational model. We then predicted e_0 for these time periods, national and subnational, and compared the subnational prediction to the observed values. Table 1 shows results for two periods out and it includes validation measures such as the bias, the mean absolute error (MAE), the continuous ranked probability score (CRPS) (Gneiting and Raftery 2007), and the coverage of the 80% probability interval (PI80). For the bias and the MAE, the smaller the better, while for the coverage of the interval, the closer to 80% the better. The CRPS provides an overall

assessment of the quality of a probabilistic forecast, including bias and variance. A better method corresponds to a larger value of CRPS.

Table 1: Out of sample validation of predicted $e_{r,c,t}^{(R)}$ for the proposed methods and other comparative methods. The “Nout” column shows how many time periods were withheld, “Nc” is the number of countries included, while “N” is the total number of data points included. The “MAE” column contains the mean absolute error, “CRPS” shows the continuous ranked probability score, and “PI80” shows the coverage of the 80% probability interval.

Nout	Method	Ncountries	N	BIAS	MAE	CRPS	PI80
2	persistence	27	874	1.457	1.463	-1.463	-
2	national	27	874	0.306	1.186	-1.473	56.2
2	shift	27	874	0.266	0.533	-0.947	89.2
2	scale	27	874	0.266	0.536	-0.950	88.8
2	AR(1)	27	874	0.265	0.534	-0.940	91.8

It can be seen that the persistence method did not perform well. The second worst was the National method, which sets the regional projections to the national values. This underperformed in terms of both MAE and bias, as well as greatly underestimating the uncertainty. The Scale, Shift and AR(1) methods performed similarly, substantially outperforming the Persistence and National methods. They slightly overpredicted uncertainty while yielding reasonable MAEs. As we have noted, the Shift and Scale methods have the undesirable property of excluding the possibility of crossovers, which the AR(1) method does not have.

4.3 Crossover Patterns

We now investigate whether the AR(1) method yields similar patterns of crossovers between regions in the same country as have been observed historically. This also gives an indication of whether the methods are representing the between-region within-country correlations well enough. To illustrate this, we randomly selected one trajectory out of the 1,000, and viewed that trajectory for all regions of a country. Figure 2 contains such plots for Brazil, Canada, Japan and the USA. It can be seen that the observed crossover pattern between regions (see lines to the left from the dotted vertical line) is similar to the predicted pattern (lines to the right of the dotted line) in each country.

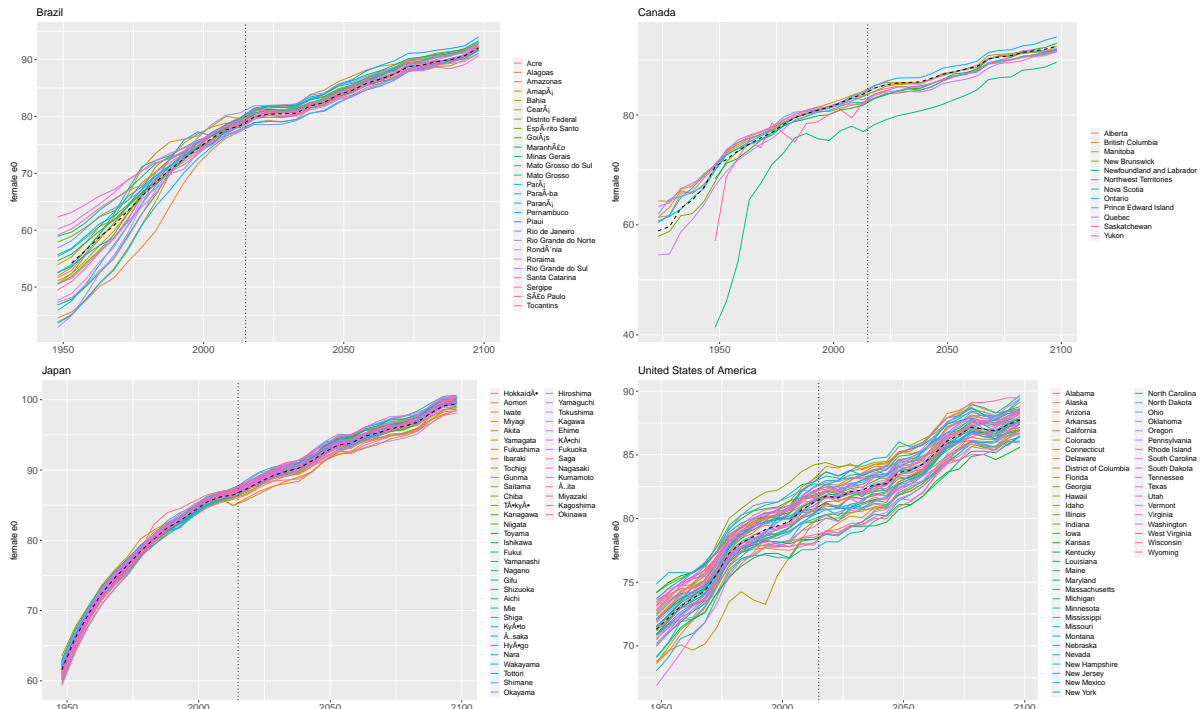


Figure 2: One-trajectory plot for Brazil, Canada, Japan and USA. Each line represents one region. Values to the left of the dotted vertical line are observed, while values to the right are results of one trajectory. National values from WPP 2017 are shown as a black dashed line.

4.4 Predictive Distribution of Subnational Mortality Rates

To project subnational age-specific mortality rates probabilistically, there are two main broad approaches one can take. One can develop a multivariate model for the age-specific rates jointly, as was done for example by Bohk-Ewald and Rau (2017). Alternatively, one can derive the distribution of the age-specific rates from the distribution of e_0 .

Here, we use the method described by Ševčíková et al. (2016), which has been used by the UN in several revisions of the WPP, including the most recent 2019 revision (United Nations 2019). In that approach, each projected trajectory of future subnational e_0 values is converted into age-specific mortality rates using a modified Lee-Carter method. Figure 3 compares results from the AR(1) method for Alaska and British Columbia in two of the future time periods, namely 2020-2025 (left) and 2080-2085 (middle and right). As expected, in the first projected time period the uncertainty is quite narrow, while 60 years later the uncertainty increases for both states and most age groups. The joint distribution of mortality rates for these two states is shown in the right panel which suggests that the probability of crossovers between regions is low across all age groups.

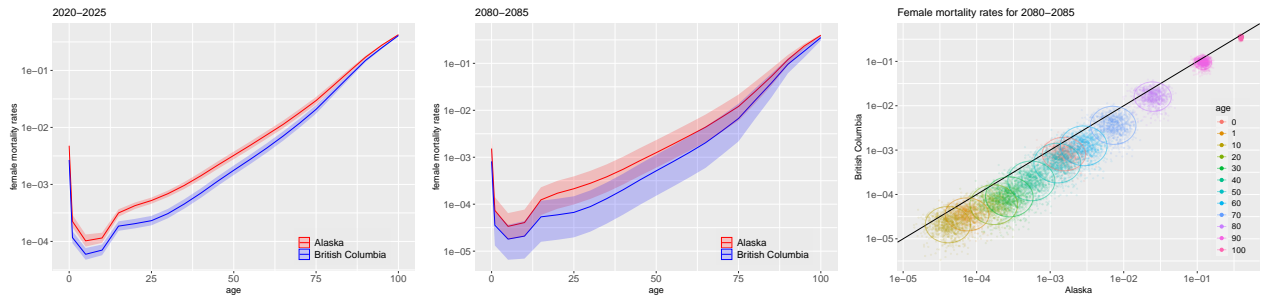


Figure 3: Predictive distribution of age specific mortality rates for Alaska and British Columbia (BC). Left and middle: Marginal distribution for Alaska (red) and BC (blue) in 2020-2025 (left) and 2080-2085 (middle). The shaded area represents the 95% probability interval with the solid lines being the medians. Right: Joint distribution between Alaska and BC in 2080-2085 for selected age groups. Each point corresponds to one trajectory. The ellipses show 95% probability intervals for each age group.

5 Discussion

We have introduced three methods for projecting subnational life expectancy that are probabilistic and build on the established Bayesian hierarchical framework for projecting national e_0 used by the United Nations. Once the national BHM has been estimated, the subnational methods are relatively simple to implement. Since trajectories from the national BHM for all countries are available for download, the additional effort required to produce subnational projections is modest. This may be an attractive feature for practitioners. All three methods can be used to generate probabilistic projections of age-specific mortality rates.

We have shown the strengths and weaknesses of the three methods. The Scale and Shift methods yield the best results in terms of the validation of the marginal predictive distribution. The biggest weakness of these two methods is that both yield perfectly parallel trajectories which is unrealistic.

The AR(1) method achieves reasonable validation of the marginal distribution, and also performs best in terms of within-country variance. It also yields crossover patterns between regions that are similar to those in the observed data, suggesting that the method reproduces the between-region correlation fairly well.

Acknowledgements: This work was supported by NICHD grant R01 HD070936.

References

- Bohk-Ewald, C. and R. Rau (2017). Probabilistic mortality forecasting with varying age-specific survival improvements. *Genus* 73.
- Gneiting, T. and A. E. Raftery (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association* 102, 359–378.
- Raftery, A. E., J. L. Chunn, P. Gerland, and H. Ševčíková (2013). Bayesian probabilistic projections of life expectancy for all countries. *Demography* 50, 777–801.
- Raftery, A. E., N. Lalic, and P. Gerland (2014). Joint probabilistic projection of female and male life expectancy. *Demographic Research* 30, 795–822.
- Raftery, A. E., N. Li, H. Ševčíková, P. Gerland, and G. K. Heilig (2012). Bayesian probabilistic population projections for all countries. *Proceedings of the National Academy of Sciences* 109, 13915–13921.

- Rayer, S., S. K. Smith, and J. Tayman (2009). Empirical prediction intervals for county population forecasts. *Population Research and Policy Review* 28, 773–793.
- Ševčíková, H., N. Li, V. Kantorová, P. Gerland, and A. E. Raftery (2016). Age-specific mortality and fertility rates for probabilistic population projections. In R. Schoen (Ed.), *Dynamic Demographic Analysis. The Springer Series on Demographic Methods and Population Analysis*, Volume 39. Springer, Cham.
- Ševčíková, H., A. Raftery, and J. Chunn (2019). *bayesLife: Bayesian Projection of Life Expectancy*. R package version 4.1-0.
- Ševčíková, H., A. E. Raftery, and P. Gerland (2018). Probabilistic projection of subnational total fertility rates. *Demographic Research* 38, 1843–1884.
- United Nations (2015). *World Population Prospects: The 2015 Revision, Probabilistic Population Projections*. New York, NY: Population Division, Dept. of Economic and Social Affairs, United Nations.
- United Nations (2017). *World Population Prospects: The 2017 Revision*. New York, NY: Population Division, Dept. of Economic and Social Affairs, United Nations.
- United Nations (2019). *World Population Prospects: The 2019 Revision*. New York, NY: Population Division, Dept. of Economic and Social Affairs, United Nations. <http://esa.un.org/unpd/wpp>.
- Wilson, T. (2018). Evaluation of simple methods for regional mortality forecasts. *Genus* 74(14), 1–22.