Bayesian Probabilistic Projections of Mortality

Jennifer L. Chunn, Adrian E. Raftery, and Patrick Gerland *

March 12, 2010

*Author's affiliations: J. L. Chunn and A. E. Raftery, University of Washington, Center for Statistics and the Social Sciences, Seattle, Washington, USA. P. Gerland, United Nations Population Division, Population Estimates and Projection Section, New York City, New York, USA. Correspondence to Jennifer L. Chunn, Center for Statistics and the Social Sciences, Box 354320, University of Washington, Seattle, WA 98195-4320. Email: chunn@u.washington.edu. The project described was supported by Grant Number 1 R01 HD054511 01 A1 from the National Institute of Child Health and Human Development. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Institute of Child Health and Human Development. Also, the views expressed in this paper are those of the authors and do not necessarily reflect the views of the United Nations. Its contents have not been formally edited and cleared by the United Nations. The designations employed and the presentation of material in this paper do not imply the expression of any opinion whatsoever on the part of the United Nations concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

Contents

List of Figures

List of Tables

1 Introduction

Every two years, the United Nations Population Division (UN) publishes the *World Population Prospects* (WPP), which include projections for populations of over 200 countries through the next 50 years. The UN accounts for uncertainty in population projections by projecting population size with total fertility rates that are higher and lower than those assumed in the main projection. However, UN projections do not account for any uncertainty in mortality projections. For comparability, a common forecasting method should be implemented across countries. We propose a probabilistic projection model of life expectancy at birth where the expected future increases in life expectancy are estimated with a Bayesian hierarchical model.

The most familiar practice for forecasting mortality is the Lee-Carter method (Lee & Carter, 1992). The Lee-Carter method produces independent country-specific forecasts based on log-linear fixed age effects and additive normally distributed homoskedastic error terms over time. When forecasting a group of countries simultaneously, a common age parameter is fixed to ensure consistent forecasts of multiple countries (Li & Lee, 2005). The Lee-Carter method has been shown to perform well (see for example Booth et al., 2005 and Bell, 1997), however, these approaches rely on the availability of age-specific death rates for at least three time periods (Li et al., 2004), which may not be available for most developing countries. In addition to data availability issues, using data for high-income countries from 1955-96, White (2002) found linear changes in life expectancy fit the observed data better than assuming linearity in log age-specific rates, which suggests projections may be better made on the life expectancy scale.

Lutz and colleagues at the International Institute for Applied Systems Analysis (IIASA)

addressed data limitation by aggregating countries into regions and forecasting regional life expectancy based on expert-based probabilistic projections (Lutz et al., 2004). Like IIASA, we project life expectancy at birth, however, we propose a random walk with non-constant drift. We model the non-constant drift using a Bayesian hierarchical model. This allows country-specific projections to be made while borrowing information from past trends of other countries.

In this paper, we develop a one-sex model for males and discuss potential extensions to a two-sex model in the closing discussion section. Because of the differences between countries in empirical data availability for the estimation of life expectancy, we use readily available expert-based estimates in this initial time-series analysis. We use male life expectancy at birth, e(0), estimates from the UN World Population Prospects (WPP) 2008 Revision from 1950 through 2010 (United Nations, 2009). Because of the significant impact of the HIV/AIDS epidemic on mortality rates, our analysis focus on countries without a generalized HIV/AIDS epidemic.

In this paper, we first briefly discuss the data, methodology used by the UN for projections and develop our proposed model, which is a natural extension of the UN's current practices. We assess the predictive power of our model via cross-validation using data from 1950-1995 to project life expectancy n 1995-2005. Then we present 90 year projections of three countries with different current life expectancies, Madagascar, Latvia, and Japan, and each with different forecasting difficulties. Life expectancy in Madagascar is amongst the lowest quartile of all countries while Japan is amongst the highest. Unlike both Madagascar and Japan, Latvia is currently in a mortality crisis and has not consistently seen improvements in life expectancy for the past few decades. For each country, we also present out-of-sample projections for 1995-2005. Lastly, we discuss aggregation of country-specific projections. Comparisons are made with the recently updated regional projections for South Asia by IIASA.

2 Data

Life expectancy at birth is a summary indicator for all age-specific mortality rates; as such, the estimation of it for over 200 countries is an arduous task. Infant and child mortality data collection and estimation is closely monitored by the international community.[1] Unfortunately, this is not the case for adult mortality. According to the 2005 UN World Mortality Report (United Nations, 2006), since 1990, only 56% of 196 countries have "reliable" or "fairly reliable" vital statistics for adult mortality. Thirty-five percent of countries have deficient or non-existent vital registration systems but have alternative sources (e.g., household death from censuses, survivorship data) to estimate adult mortality which may not always be reliable. Lastly, 7% of countries are lacking recent data for the estimation of adult mortality completely.

This disparity in reliable data availability is not homogenous across populations. Of the 50 countries in Asia, 56% have "reliable" or "fairly reliable" vital statistics, whereas 95% of Europe and North America have reliable statistics. This number decreases dramatically in Africa where only five of the 54 countries (9%) maintain "reliable" or "fairly reliable" vital statistics.

Because of the inequities in data reliability and availability, for this analysis of the projection of mortality, we used the life expectancy at birth time series from the World Population Prospects 2008 Revision (WPP) produced by the United Nations Population Division (United Nations, 2009). WPP estimates are an expert culmination of the disparate data and methodological machinery available.

The HIV/AIDS epidemic has been a major source of mortality in the last 20 years. As such, we exclude countries with a generalized HIV/AIDS epidemic.[2] A country is defined as

[1] See for example the Inter-agency Group for Child Mortality Estimation with members from the United Nations Children's Fund, World Health Organization, The World Bank and United Nations Population Division

[2] Classification of countries with a generalized HIV/AIDS epidemic was based on HIV/AIDS fact sheets

having a generalized epidemic when: (a) HIV is established in the general population; (b) the epidemic could be sustained via sexual networking in the general population independent of sub-populations at higher risk for infection; (c) HIV prevalence is consistently over 1%. A total of 158 countries were included in the model.

3 Methodology

3.1 Model

Currently, the UN projects life expectancy deterministically. The life-expectancy ($l_{c,t+1}$) for country, c, in the next quinquennial period, $t+1$, is estimated to be the life expectancy in the current time period ($l_{c,t}$) plus the expected gains in life expectancy ($g(l_{c,t})$). Observed five-year gains in life expectancy for 158 countries from 1950 to 2005 are plotted in Figure 1. This figure highlights the non-constant rate of change in life expectancy. To capture this, the UN has developed models that represents the decline in mortality by fitting a double-logistic function of current life expectancy. The five deterministic UN models are displayed in Figure 1(b), where models vary by pace of gains in life expectancy.

published jointly by WHO, UNAID and UNICEF (2008).

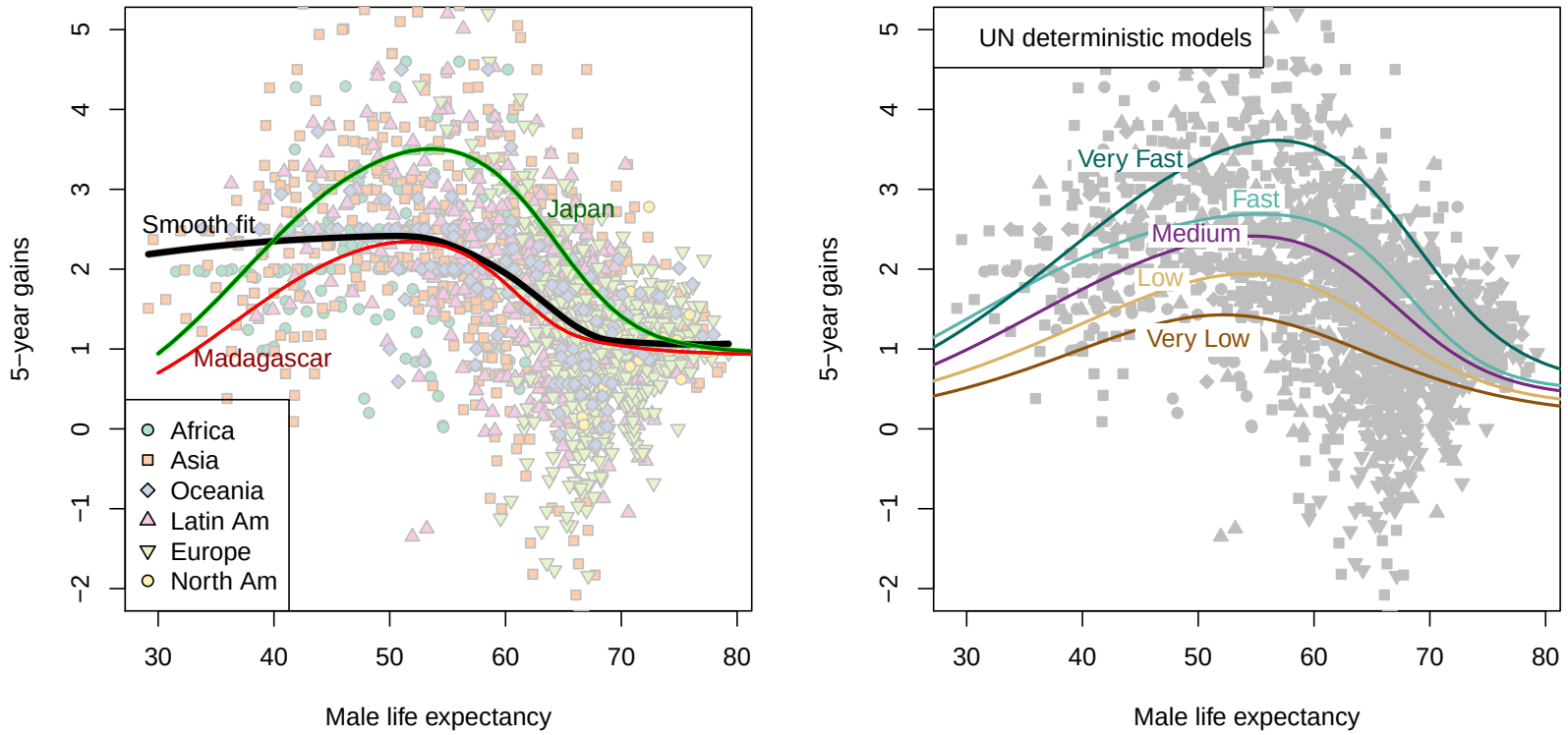


Figure 1: Observed five-year gain in life expectancy within a country. UN estimates for 158 countries from 1950 to 2005 are included in this figure (n=1738). Each point represents the observed five-year gain in life expectancy within a country. The black line is a locally-weighted polynomial regression (lowess) of the observations, which highlights the non-constant rate of gains in life expectancy. (Note, 31 observations (1.8%) are outside the range of the plot and not depicted, but where include in the local regression.) Included in the left plot are the fitted posterior median double-logistic functions for Japan and Madagascar from our model. The UN deterministic models are included in the right plot.

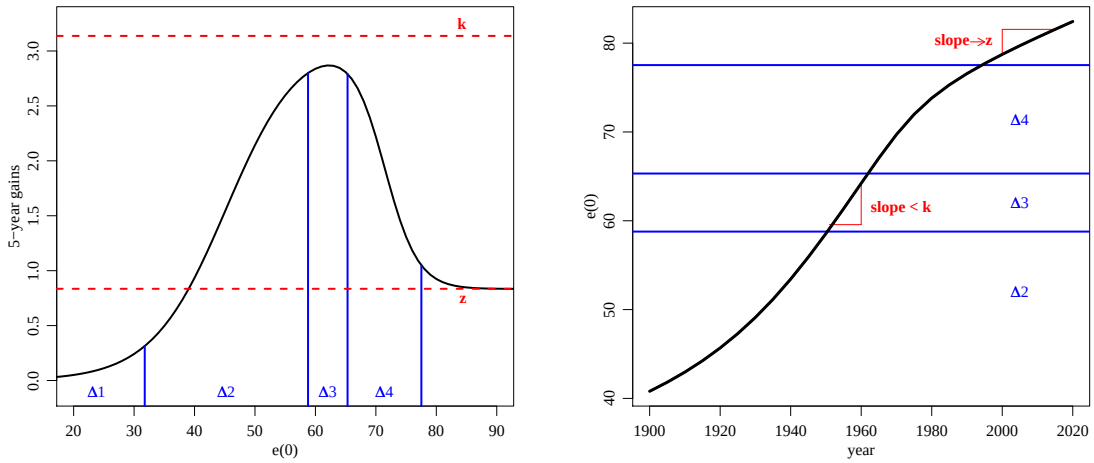


Figure 2: Illustration of the double-logistic function (based on a curve from Japan's posterior distribution). The left plot illustrates the double-logistic function of 5-year gains in life expectancy. The right plot is a time-series of life expectancy, e(0), with gains modeled according to the double-logistic function

The double-logistic function has six parameters, as is illustrated in Figure 2, four identifying intervals of life expectancy when the rate of life expectancy gains is changing, one describing the approximate maximum gain in life expectancy, and the last describing the asymptotic rate of gains as life expectancy increases (Meyer, 1994). For each country, a UN analyst chooses one of five prescribed choices of the six parameters[3] by assessing the recently observed pace of mortality decline (United Nations, 2009). In the UN approach, a constant rate is set once the gains in life expectancy have reached a preset low level. Although there is no evidence of an upper limit to life expectancy (Oeppen & Vaupel, 2002), setting future life expectancy gains to be constant assumes life expectancy in all countries will continue to rise at the same rate and not stay the same or decline.

The double-logistic function is the sum of two 3-parameter logistic growth pulses. Demo-

[3]Using data from different percentiles, the UN developed five models for each gender to describe different rates of increase in life expectancy: very fast (based on upper 10^{th} percentile), fast (upper 25^{th} percentile), medium (based on median), slow (lower 25^{th}), and very slow (lower 25^{th} percentile).

graphic transition theory suggests slow increase in life expectancy then significant increases in life expectancy as a country enters the demographic transition. This increase in life expectancy is not constant over time. Rapid gains in life expectancy are a result of improvements in infant and child mortality. However, gains slow as mortality improvements shift to older ages.

To summarize, the UN method to estimate the life expectancy in the next time period is given by

$$l_{c,t+1} = l_{c,t} + g(l_{c,t}). \tag{1}$$

The expected 5-year gains in life expectancy is a function of the current level of life expectancy as determined by a UN analyst chosen parameterization of the double-logistic function

$$\begin{aligned} g(l_{c,t}|\boldsymbol{\theta}^c) &= \frac{k^c}{1+exp(-\frac{log(9^2)}{\Delta_2^c}(l_{ct}-\Delta_1^c-0.5\Delta_2^c))} + \frac{z^c-k^c}{1+exp(-\frac{log(9^2)}{\Delta_4^c}(l_{ct}-\sum_i \Delta_i^c+0.5\Delta_4^c))} \\ \boldsymbol{\theta}^c &\in (\boldsymbol{\theta}^{\text{Very Slow}}, \boldsymbol{\theta}^{\text{Slow}}, \boldsymbol{\theta}^{\text{Medium}}, \boldsymbol{\theta}^{\text{Fast}}, \boldsymbol{\theta}^{\text{Very Fast}}) \\ \boldsymbol{\theta}^c &= (\Delta_1^c, \Delta_2^c, \Delta_3^c, \Delta_4^c, k^c, z^c). \end{aligned}$$

A natural extension to the account for the uncertainty in the UN methodology would be to model the underlying generating mechanism as a random walk with drift where the drift term is given by the double-logistic function. This means that life expectancy in the next time period is equal to the UN estimate plus a random perturbation ($\epsilon_{c,t+1}$):

$$l_{c,t+1} = l_{c,t} + g(l_{c,t}|\boldsymbol{\theta}^{(UN)}) + \epsilon_{c,t+1}. \tag{2}$$

This simple extension accounts for uncertainty around the UN analyst chosen parametric double-logistic function, yet, it does not account for the uncertainty associated with the

chosen set of double-logistic parameters. We use the UN expert knowledge by assuming the rate of gains in life expectancy follow this flexible double-logistic function. However, we do not assume that countries will follow a specific double-logistic function with preset parameters. We propose modeling the drift term as a non-linear Bayesian hierarchical model, which allows country-specific double-logistic parameters to be fit and pool information about the rate of gains across countries conditional on the level of life expectancy at birth. A Bayesian-based approach allows the estimation of country-specific probabilistic distributions of gains in life expectancy. Our model is given by

$$l_{c,t+1} = l_{c,t} + g(l_{c,t}|\boldsymbol{\theta}^{(c)}) + \epsilon_{c,t+1} \tag{3}$$

where,

$$\begin{aligned}
g(l_{c,t}|\boldsymbol{\theta}^c) &= \text{Double-Logistic function with parameters } \boldsymbol{\theta}^c \\
\boldsymbol{\theta}^c &= (\Delta_1^c, \Delta_2^c, \Delta_3^c, \Delta_4^c, k^c, z^c) \\
\Delta_i^c|\sigma_{\Delta_i} &\sim_{ind} \text{Normal}_{[0,100]}(\Delta_i, \sigma^2_{\Delta_i}) \\
k^c|\sigma_k &\sim_{ind} \text{Normal}_{[0,10]}(k, \sigma_k^2) \\
z^c|\sigma_z &\sim_{ind} \text{Normal}_{[0,1.15]}(z, \sigma_z^2)
\end{aligned}$$

We pool information about the rates of gains across countries by assuming each set of country-specific double-logistic parameters are randomly sampled from a common normal distribution around a set of world double-logistic parameters. The normal distribution is truncated such that all of the double-logistic parameters are positive. It is easily evident why the first five parameters should be positive since they are intervals of life expectancy and the maximum gains, respectively. The upper truncation points for the first five parameters were arbitrarily chosen to be large. Sensitivity analysis of the values showed results were not

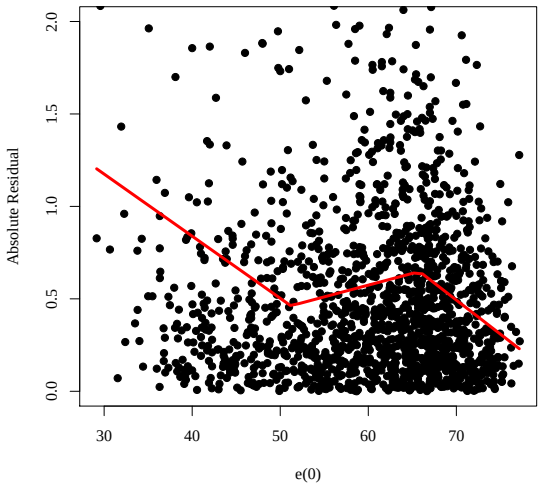


Figure 3: Homoskedastic model absolute residuals across life expectancy with fitted regression spline, which highlights the heteroskedasticity. (Note, 44 (2.8%) residuals are outside the range of the plot, but were included in the fitted spline.)

affected by changes to what is deemed "large".

In their seminal paper, Oeppen & Vaupel (2002) found a strong linear trend in the "best practices" life expectancy (i.e., highest life expectancy of a given year) from mid-19th century through 2000. By assuming the sixth parameter, z, is non-negative, we are assuming that on average, life expectancy will continue to increase, which is consistent with the findings of Oeppen & Vaupel (2002). For a fitted linear regression of male life expectancy, they reported $r^2 = 0.980$ and a slope of 0.222 (increase per year), which equates to a 1.11 year increase in life expectancy per quinquennial. Because this is the rate of increase for "best practices" countries, we assume that the asymptotic rate of increase for any given country will not exceed this rate. To relax this assumption, we assume no country will asymptotically out pace the 99.9^{th} percentile rate, 1.15.

Figure 3 shows the absolute residuals when we assume homoskedasticity around the double-logistic function. As is evident from the regression spline of these residuals, the obser-

vations are not scattered around the double-logistic function in an equal pattern. Instead, the distribution around the function decreases as life expectancy increases. Our model addresses this heteroscedasticity by assuming the standard deviation of the random perturbations are proportional (ω) to the absolute residual regression spline of life expectancy ($f(l_{c,t})$), where knots are determined by the posterior mean of $\Delta_1 + \Delta_2 + \Delta_3$ and $\Delta_1 + \Delta_2 + \Delta_3 + \Delta_4$, life expectancies when the rate of increase changes. We model the stochastic error term around the double-logistic function as

$$\epsilon_{ct} \sim_{ind} N(0, (\omega \times f(l_{c,t-1}))^2) \tag{4}$$

For the stochastic proportionality constant, ω, a diffuse prior was specified uniformally from zero to ten, where the upper limit is specified to be slightly larger than the ratio of observed standard deviation in life expectancy gains to the smallest functional value of $f(\cdot)$. Inverse-gamma priors were specified for variance parameters with prior variance estimated by the empirical average squared deviations from the UN medium-pace model, $(15.6^2, 23.5^2, 14.5^2, 14.7^2, 3.5^2, 0.6^2)$. Priors of the world-level parameters were assumed to be sampled from a truncated normal density with normal means and variances set to the UN medium-pace model and variance of the UN models, $(15.77, 40.97, 0.21, 19.82, 2.93, 0.40)$ and $(3.56^2, 3.93^2, 3.96^2, 3.80^2, 0.99^2, 0.16^2)$, respectively. All prior distributions were much more diffuse than posterior distributions.

3.2 Parameter estimation

The posterior distribution of the parameters was sampled via a Gibbs sampler where each conditional posterior distribution was explored using slice sampling Markov Chain Monte Carlo (MCMC) methods (Neal, 2003). Models were run using the package R2WinBUGS (2005) in R 2.7 (2009) to access WinBugs 1.4 (2000).

Although we cannot guarantee the MCMC has reached its stationary distribution, we can rely on a few ad hoc diagnostic tools to highlight situations where the chain does not appear to have reached stationarity. Visual inspection of trace plots for trending in the sample space can suggest lack of convergence. In addition to convergence to the stationary distribution, it is important that the chain has mixed well, that is, has explored the entire posterior distribution. High auto-correlation of the chain suggests the chain is mixing slowly and hence long chains are necessary. Because of the nature of Gibbs sampling whereby one parameter is updated conditional on the current state of the other parameters, the resulting chains typically mix slowly when the parameters are correlated.

More formal diagnostic tools have been developed to help identify situations of non-convergence. The Raftery-Lewis diagnostic (Raftery & Lewis, 1992) is a rough approximation of the length a chain needs to be run in order to estimate a desired percentile within a certain degree of accuracy. For an initial chain of 10,000 iterations, Raftery-Lewis diagnostics suggest that for the vast majority of parameters, a chain length of 25,000 would be appropriate and a length of 100,000 for all parameters well-defined in the likelihood.

The Gelman-Rubin (G-R) statistic (Gelman & Rubin, 1992) uses the idea that chains starting from different initial values should act similarly once stationarity has been reached. As such, the G-R statistic compares the within chain variation to the between chain variation where the statistic should converge to 1. By iteration 50,000, the statistic converged to 1, which suggests the chain has reached convergence.

For each run of the model, three chains of length 100,000 were run with a burn-in of 10,000. Chains were thinned to control the file size. As suggested by the Raftery-Lewis and Brooks-Gelman-Rubin diagnostics, the trace plots of the parameters suggest the chains were well mixed and had converged.

4 Model Validation: Assessing the predictive ability

To assess the validity of our forecasts, we evaluated the calibration and sharpness of our predictions (see Gneiting et al. (2007) for full discussion of diagnostics). Calibration compares our predictive distributions with the actual observations, while sharpness refers to the concentration of predictive distribution. The ideal projections would be the sharpest (i.e., narrowest prediction intervals) without sacrificing calibration (i.e., accurate predictions). Cross-validation was performed by fitting our model to data from 1950 through 1995 (n=1,422) and forecasting life expectancy for males from 1995 to 2005, resulting in 316 cross-validation points.

When assessing the predictive ability of our model, we examined numerical measures of calibration via the coverage of our prediction intervals, root mean squared error (rMSE), mean absolute error (MAE) and the mean standardized absolute predictive error (SAPE). Standardized absolute predictive errors (a_{ct}) are defined for country, c, at time period, t, as:

$$a_{ct} = \sqrt{\frac{2}{\pi}} * \frac{|l_{ct} - \hat{l}_{ct}|}{\hat{\sigma}_{pred,ct}}.$$

That is, the SAPE is the absolute difference between the actual observed life expectancy (l_{ct}) and our median forecast ($\hat{l}_{ct}$) standardized by the standard deviation of our predictive distribution. When the model is correctly specified, the expected mean SAPE value is equal to one.

These numerical metrics are presented in Table 1. Overall, our model was well calibrated with our 95% prediction intervals capturing the actual observations 92% of the time. The nominal coverage of our 80% prediction intervals was 82%. The mean standardized absolute error (SAPE) was 1.04, which is quite close to the theoretical mean of 1. The mean absolute error (MAE) of our median predictions was 1.07 year. That is, over 10 years of predictions,

our "best guess", on average, was within 1 year of the actual observation

The predicitve ability of our model was compared with the current UN methodology. Replicating our cross-validation methodology by using WPP 2008 data through 1995, a UN analyst computed life expectancy forecasts for 1995-2005 using one of the five prescribed UN models of gains in life expectancy at birth based on levels and trends in the preceding two decades. As is seen in Table 1, the root mean square error and mean absolute error of the UN cross-validation were both larger than our cross-validation results. Our model reduced the mean absolute error by 42%.

Table 1: BHM summary measures for 10 year out-of-sample cross-validation. Cross-validation was also performed using the UN's current methodology and are included for comparability. We see that our model improved the mean absolute error by 42%.

Measure	BHM	UN cross-val
Root mean squared error (rMSE)	1.64	2.50
Std absolute prediction error (SAPE)	1.04	n/a
Mean absolute error (MAE)	1.07	1.86
Prediction Intervals		
Nominal	Actual	Mean half-length
95%	92.1%	2.54
90%	89.2%	2.13
80%	82.0%	1.66

Gneiting et al. (2007) proposed an assessment of probabilistic performance be based on maximizing sharpness subject to calibration. That is, calibration being equal, the more concentrated the predictive distribution, the better. Because the UN does not produce probabilistic projections, we were unable to compare the sharpness and calibration of our results with theirs. However, we evaluated the sharpness of our projections by examining the distribution of prediction interval lengths. For the 1995-2000 time period, the prediction interval half-lengths range from 0.7 to 1.9 years with an average half-length of 1.3 years. For the next quinquennial period, 2000-2005, the interval half-lengths increase to a range of

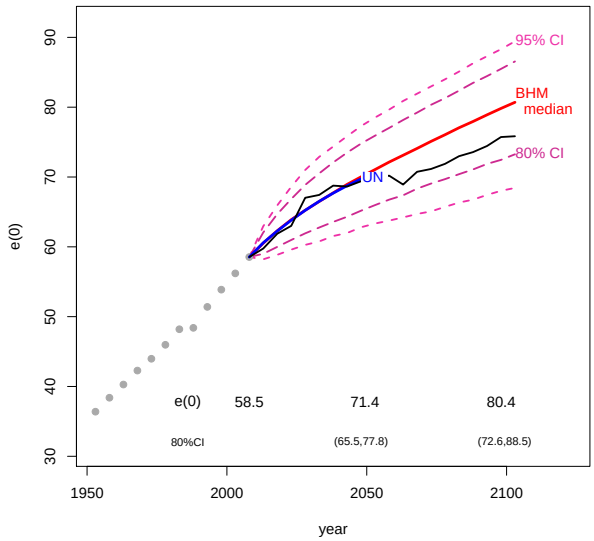
1.0 to 3.2 years with an average half-length of 1.9 years. Both life expectancy at birth and prediction interval lengths vary by region. From 1995 to 2005, Africa had the lowest life expectancy of 59.8 years with an average interval half-length of 2.2 years. With an average life expectancy of 73.5 years and interval half-length of 1.3 years, North America had both the highest life expectancy and the most narrow prediction intervals, however, this region only consists of two highly correlated countries (US and Canada).

5 Country-specific Case Studies

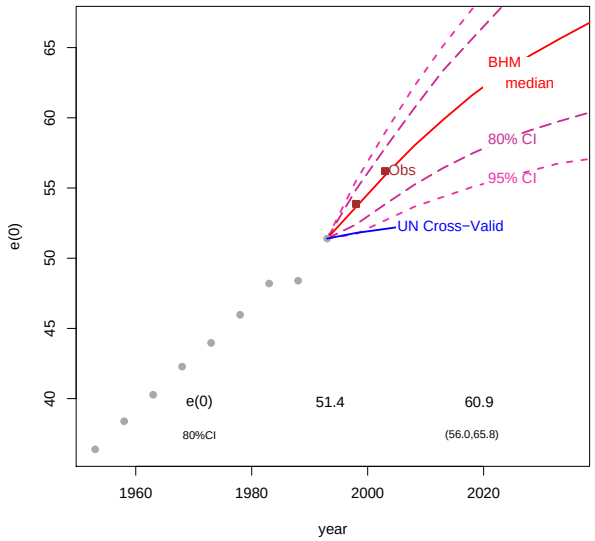
5.1 Typical country: Madagascar

Located off the coast of southern Africa, Madagascar is an island nation of over 17 million people. Currently, the UN estimates life expectancy at birth among males in Madagascar to be 58.5 years. Figure 4(a) displays projections of life expectancy starting from the 2005-2010 quinquennial. We see that our BHM median forecast is almost identical to WPP 2008 forecasts through 2050. In 2045-2050, the UN projects male life expectancy will be 69.7. We project life expectancy will be 71.4 years with an 80%CI of (65.5,77.8). Fifty years after that, in 2095-2100, we project life expectancy to reach 80.4 years with a wider 80%CI of (72.6,88.5). In Figure 4(b), we present out-of-sample projections for Madagascar with projections beginning in 1990-1995. UN observed estimates are indicated in the plot by brown squares and align along our median projection. The exclusion of two time periods results in more uncertainty in our projections for 2095-2100 with a median of 76.6 years and a 80%CI of (64.7,87.8).

Along with quantiles of the projected life expectancy distribution, in Figure 4(a) (as well as Figures 5(a) and 6(a) for Latvia and Japan), we also include a sample stochastic trajectory for Madagascar. We see that unlike the quantiles, the sample trajectory does not follow a



(a) Projections from 2005-2010.



(b) Cross-Validation projections from 1990-1995.

Figure 4: Life expectancy projections for Madagascar. The above plots include the UN projections and our median projections, with 80% and 95% CIs. The life expectancy values used in our modeling are indicated by grey circles. Figure 4(a) show projections from 2005-2010 with a sample BHM trajectory. Our median projections (BHM) are quite similar to those by the UN through 2050. Our 80%CI in 2050 is roughly ±5 years around the UN deterministic projection. Figure 4(b) shows cross-Validation projections from 1990-1995. Observed life expectancies from 1995-2005 are shown as squares, which we see were closely projected by our model.

smooth path. For this trajectory the mean absolute deviation from the median is that of the median mean absolute deviation. That is, it is a trajectory with typical deviation from the projected median.[4]

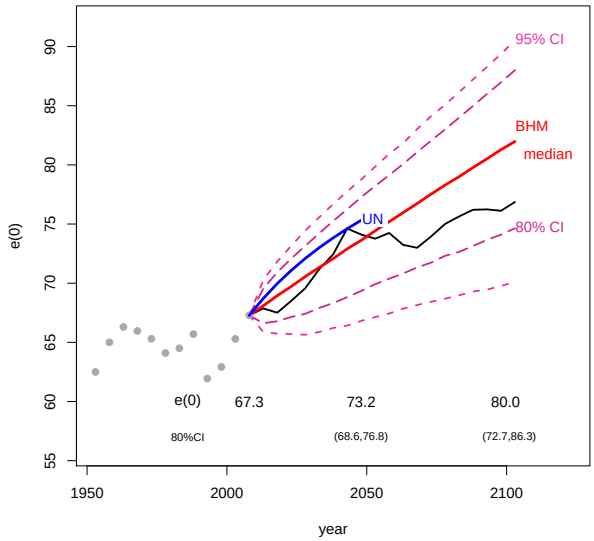
5.2 Mortality Crisis: Latvia

Figure 5 is a plot of estimated and projected life expectancy for Latvia. Male life expectancy in Latvia was increasing from 62.5 years in 1950 to 66.3 years in 1965. However, in the subsequent 15 years, male life expectancy in Latvia *decreased* by 2.2 years to 64.1 years.

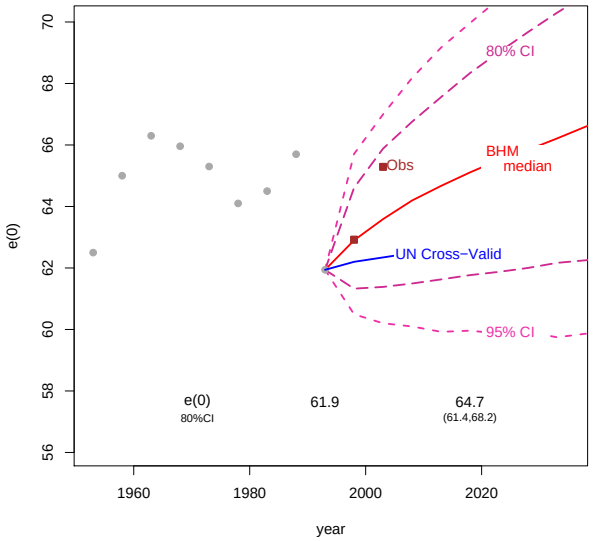
[4]This illustration of a typical stochastic trajectory was suggested by Joel Cohen (personal communication, December,2009).

Life expectancy increased again until a 3.8 year decline between the 1985-1990 and 1990-1995 quinquennial periods was observed. In the past 10 years, it appears Latvia has recovered from the decline in life expectancy. In Figure 5(a), both the UN and our median projections predict a continuous increase in life expectancy, however, ours increase at a more conservative rate.

As is evident in Figure 5(b), our 80% prediction intervals capture the "true" observed estimates of life expectancy from 1995 to 2005. For the first time period, from 1995-2000, the upper bound of our 80% prediction interval was 64.4 years. Yet, the lower bound of our 80% prediction interval, 61.1 years, actually predicts that life expectancy may continue to decrease. In fact, our prediction interval allows for the possibility of life expectancy not increasing for the subsequent 50 years, which highlights the uncertainty of the previous time-series.



(a) Ninety year projections starting from 2005-2010.



(b) Cross-Validation projections from 1990-1995.

Figure 5: Life expectancy projections for Latvia. The above plots include the UN projections and BHM median projections, with 80% and 95% CIs. The life expectancy values used in our modeling are indicated by grey circles. Figure 5(a) contains ninety year projections starting from 2005-2010. A typical stochastic trajectory is included illustrating the non-smoothness of individual projections. We see that the UN projections through 2050 closely coincide with our median projection. However, given the volatility in past life expectancy, the lower bound of our CIs account for the potential for e(0) to decline again in the future. Figure 5(b) contains cross-Validation projections from 1990-1995. Observed life expectancies from 1995-2005 are shown as squares, which we see were closely projected by our model. By 1995, Latvia had not yet recovered from its mortality crisis; our projection intervals reflect the uncertainty of a full recovery.

5.3 Leading country: Japan

One of the difficulties with projecting mortality is accurately projecting the country with the lowest mortality and, hence, the highest life expectancy. Historically, "pessimists" believed that life expectancy could not keep rising at the historic rates and assumed there must be a "ceiling" to life expectancy for humans (see for example, Fries, 1980,Olshansky et al., 1990, Olshansky & Dsesquelles, 2001, Olshansky et al., 2002, Olshansky et al., 2005). There are others, "optimists", who believe life expectancy will continue to increase without limit (see for example, Oeppen & Vaupel, 2002, Tuljapurkar et al., 2000, Tuljapurkar, 2005). However, past estimates of the "maximum life expectancy" have continually been surpassed (Oeppen & Vaupel, 2002) and old-age mortality rates continue to decline (Vaupel et al., 1998). In fact, Oeppen & Vaupel (2002) presented strong evidence that the world's highest, or "best practices", life expectancy at birth has increased linearly across time and show no signs of leveling off. They estimated that the "best practices" life expectancy for males has increased at a rate of 0.222 per year. Although Japan does not have the highest male life expectancy in the current quinquennial (that title has belonged to Iceland since 2000), it is the country with the highest overall life expectancy and has been since 1980. Figure 6(a) is a plot of male life expectancy in Japan. Also included in the plot is what the trajectory would be if male life expectancy in Japan increased at the "best practices" rate of 1.11 per quinquennial. Vallin & Meslé (2009) updated and expanded the data time period (from 1840-2000 to 1750-2005) for "best practices" life expectancy. They found a segmented line fit the extended time frame better, with the most recent segment (1960-2005) still with a strong positive slope (1.13 years per quinquennial for women) and conclude the Oeppen and Vaupel line may be too optimistic for the long-term future. Our median projection are more conservative than the Oeppen-Vaupel "best practices" linear projection, suggesting the potential over-optimism of the "best-practices" lines. However, the "best practices" trajectory is just within the upper

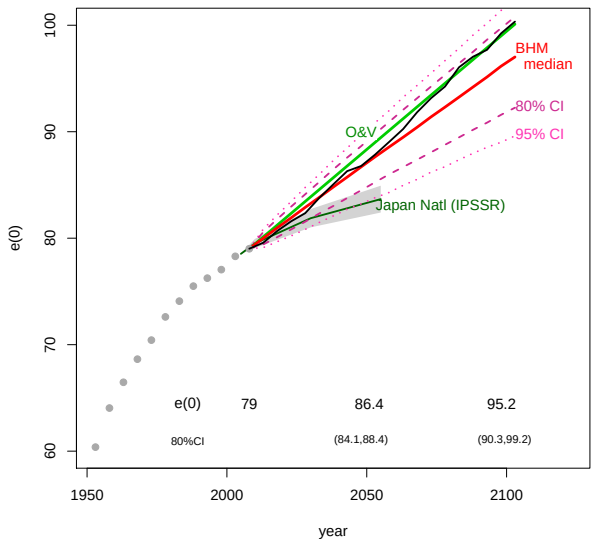
bound of our 80%CI.

Bongaarts (2006) also found that the Oeppen-Vaupel "best-practices" rate to be overly optimistic. By decomposing mortality into juvenile, background, and senescent mortality, Bongaarts (2006) observed that historically large gains in life expectancy were due to declines in juvenile mortality. Then as juvenile mortality reached low levels, the rate of gains in life expectancy diminished. This decomposition supports the rationale for modeling the gains in life expectancy with a double-logistic function where there are periods of high gains in life expectancy (i.e., when juvenile mortality is declining) followed by a leveling of gains (i.e., when the gains in life expectancy are due to incremental declines in senescent mortality). Bongaarts (2006) found that senescent life expectancy in countries with low mortality, on average, increased at a rate of 0.15 years per year, or 0.75 years per quinquennial. The average asymptotic rate of gains estimated in our model was 0.17 years per year (0.84 years per quinquennial), which is much closer to the Bongaarts (2006) projected gains in senescent life expectancy then the "best practices" rate of increase.

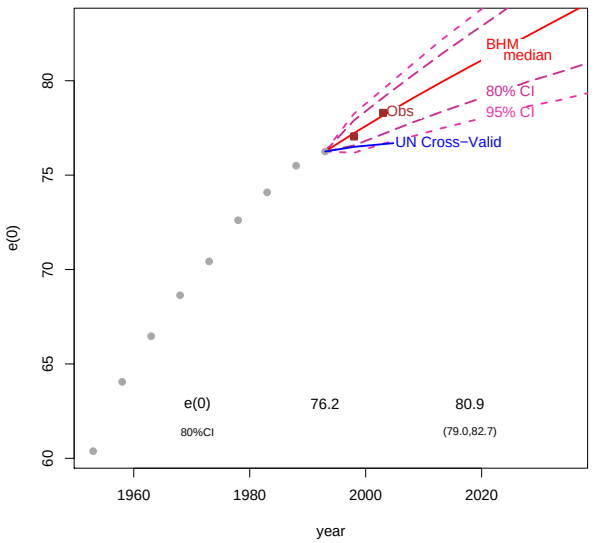
Recently, the Japanese official projections made by the National Institute of Population and Social Security Research (IPSSR) have extended the traditional Lee-Carter method to better estimate mortality at higher ages. The original Lee-Carter model estimates age-specific mortality rates for five year intervals with the last age range aggregating those 85 and older. IPSSR now use the shifting logistic model (Bongaarts, 2005) to account for continued increase in life expectancy in Japan. IPSSR projections (low/medium/high rates of mortality decline variants, with the medium variant being equivalent to the UN projections) (Kaneko et al., 2008) are included in Figure 6(a). The IPSSR are more conservative and project an earlier leveling off of life expectancy than our projections, but are still within our prediction intervals.

When looking at out-of-sample projections in Figure 6(b), which begin in 1990-1995, we see the UN cross-validation projections suggest an immediate leveling off of life expectancy.

Yet, the observed life expectancy in 1995-2005 did not level off and instead continued to increase.



(a) Projections beginning in 2005-2010



(b) Cross-Validation projections from 1990-1995.

Figure 6: Life expectancy projections for Japan. The above plots include the UN projections and BHM median projections, with 80% and 95% CIs. The life expectancy values used in our modeling are indicated by grey circles. Figure 6(a) contains projections from 2005-2010 with a sample trajectory. National Institute of Population and Social Security Research (IPSS) medium variant are the same as the UN projections with uncertainty bounds indicated in the shaded region. We include a trajectory with a constant increase of 0.222 per year, as found by Oeppen and Vauppel (O & V) (2002) among the "best practices" country. A typical stochastic trajectory is also included to illustrate the non-smoothness of individual BHM projections. We see that our median projections are between the "best practices" line and the official projections. Figure 6(b) depicts cross-Validation projections from 1990-1995. Observed life expectancies from 1995-2005 are shown as squares. Our model (BHM) closely predicted the continued rise in e(0), whereas the UN cross-validation would have underestimated.

6 Aggregation to Regional Projections: South Asia

As previously mentioned, researchers at the International Institute for Applied Systems Analysis (IIASA) (Lutz et al., 2004) produced regional probabilistic prediction intervals for

life expectancy using Delphi-type methods. Linear paths were then drawn from a normal distribution to produce probabilistic predictive distributions. A strength of this method is that it uses demographic knowledge as an input whereas traditional time-series methods only rely on past trends.

Country-specific projections allow regional projections to be made regardless of how the region is defined. To compare our projections with those of IIASA for South Asia, we aggregated UN estimates and projections and our projections to be proportional to the regional male average populations from 2005 to 2010.[5] The countries included in the IIASA defined region of South Asia (with population percentage) are: India (75.1%), Pakistin (10.5%), Bangladesh (9.9%), Nepal (1.7%), Afghanistan (1.5%), Sri Lanka (1.3%), Bhutan (0.04%), Maldives (0.02%).

The model we presented above assumes that the random distortions in gains in life expectancy are independent across countries. Previous work (Alho, 2008) has suggested that cross-country correlations are non-zero and should be modeled as such. Within South Asia, in the past 60 years, life expectancy was indeed correlated where the correlations are as follows:

$$\begin{pmatrix}
\text{Afghanistan} & 0.91 & 0.91 & 0.99 & 0.98 & 0.94 & 0.98 & 0.99 \\
0.91 & \text{Bangladesh} & 0.99 & 0.92 & 0.96 & 0.99 & 0.97 & 0.89 \\
0.91 & 0.99 & \text{Bhutan} & 0.92 & 0.96 & 0.99 & 0.96 & 0.89 \\
0.99 & 0.92 & 0.92 & \text{India} & 0.99 & 0.95 & 0.99 & 0.99 \\
0.98 & 0.96 & 0.96 & 0.99 & \text{Maldives} & 0.98 & 0.99 & 0.98 \\
0.94 & 0.99 & 0.99 & 0.95 & 0.98 & \text{Nepal} & 0.98 & 0.92 \\
0.98 & 0.97 & 0.96 & 0.99 & 0.99 & 0.98 & \text{Pakistan} & 0.97 \\
0.99 & 0.89 & 0.89 & 0.99 & 0.98 & 0.92 & 0.97 & \text{Sri Lanka}
\end{pmatrix}$$

[5]Regional aggregation by weighted average using total male population is an approximation to the true regional aggregation. The true computation is more complex and requires the survival ratios by age and sex weighted by their respective population by age and sex. In addition, future weights depend on future fertility by country, and therefore future weights will change over time. Nevertheless, from a quick comparison, we found for 1950-2050 on South-Central Asia life expectancy differences of at most +0.3 years for males.

Although life expectancy was highly correlated within South Asia, this does not necessarily mean the random distortions in our model need to be projected with such a strong association. The hierarchical modeling of the change in life expectancy inherently allows for between country correlation in life expectancy. Instead, we are interested in the residual correlations in life expectancies gains, ρ_{c_i,c_j}, between countries c_i and c_j. Below is the matrix of posterior predictive expectations of these residual correlations, whose values are much closer to 0 than the correlation matrix of life expectancy above.

$$\left(\hat{\rho}_{c_i,c_j}\right) = \begin{pmatrix} \text{Afghanistan} & -0.28 & -0.23 & 0.09 & 0 & -0.07 & -0.08 & 0.19 \\ -0.28 & \text{Bangladesh} & 0.05 & -0.03 & 0.03 & 0 & 0.05 & -0.19 \\ -0.23 & 0.05 & \text{Bhutan} & -0.03 & 0.05 & 0.09 & 0.07 & 0.2 \\ 0.09 & -0.03 & -0.03 & \text{India} & 0.16 & 0.03 & 0.03 & 0 \\ 0 & 0.03 & 0.05 & 0.16 & \text{Maldives} & 0.24 & 0.25 & 0.34 \\ -0.07 & 0 & 0.09 & 0.03 & 0.24 & \text{Nepal} & 0.08 & 0.05 \\ -0.08 & 0.05 & 0.07 & 0.03 & 0.25 & 0.08 & \text{Pakistan} & -0.01 \\ 0.19 & -0.19 & 0.2 & 0 & 0.34 & 0.05 & -0.01 & \text{Sri Lanka} \end{pmatrix}$$

Country-specific projections were then made by sampling the vector of random distortions, $(\delta_{c_i,t})$, from a multivariate normal distribution whose covariance matrix incorporates the residual correlation matrix, $\left(\hat{\rho}_{c_i,c_j}\right)$. The aggregated projections for the South Asian region are presented in Figure 7. The 2007 IIASA projections available on their website (Lutz et al., 2008) are also depicted. We found that our median projections were very consistent with IIASA median projections, yet, our intervals are more sharp, ranging from 36-72% more narrow than those of IIASA.

In addition to improved precision, our approach is a random trajectory that does not assume temporal correlation across the length of the projection. As such, our sample trajectories can fluctuate around the mean life expectancy. In contrast, IIASA samples random linear trajectories. In this case, a trajectory that is higher (lower) than the mean life expectancy, will remain higher (lower) than the mean life expectancy.

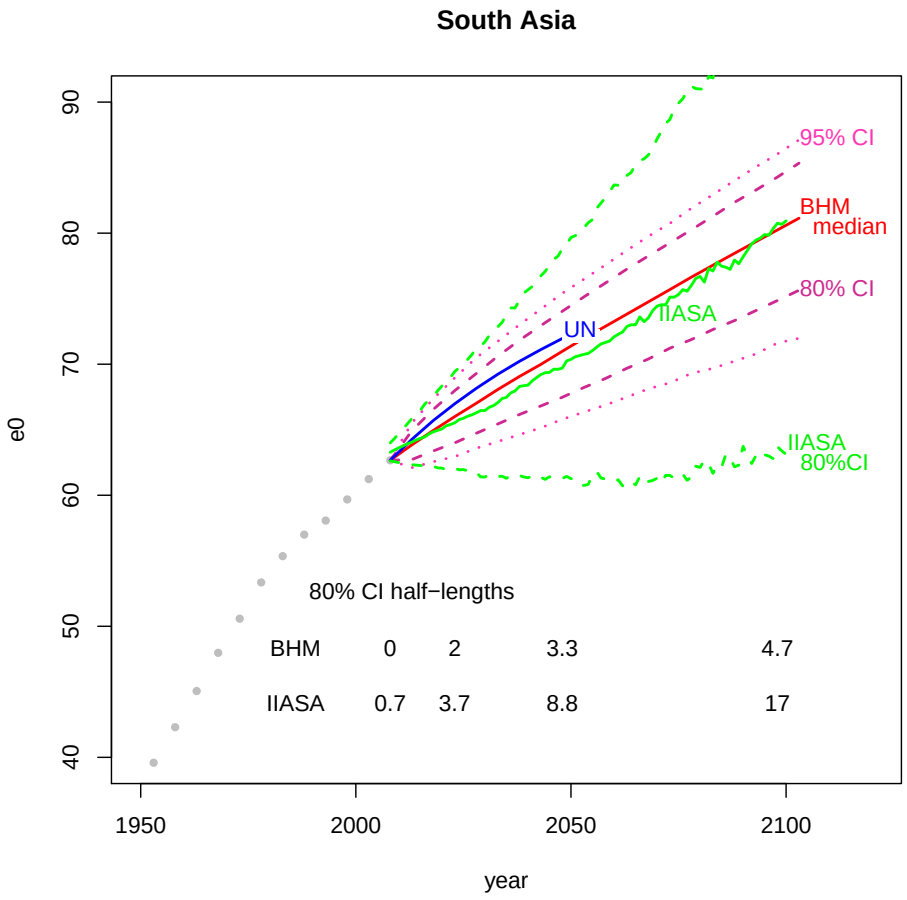


Figure 7: Life expectancy projections for South Asia (IIASA-defined) for our BHM model, IIASA and the UN. The median projections for BHM and IIASA are similar with each other, but the IIASA 80% intervals are wider than the BHM 80% interval.

7 Discussion

Much research has been done on the forecasting of mortality (see Booth, 2006 for a detailed review). However, efforts have focused on developed countries where reliable age-specific data are available. As previously discussed, the most ubiquitous time-series method for forecasting age-specific mortality rates is the Lee-Carter method and its various parallels (e.g., Renshaw & Haberman, 2006), generalizations (e.g., de Jong & Tickle, 2006; Hyndman & Ullah, 2007; Pedroza, 2006; Koissi et al., 2006; Brouhns et al., 2002) and extensions (e.g., Li & Lee, 2005; Li et al., 2004; Ishii, 2008). Furthermore, Lee & Miller (2001) found the forecasts of the Lee-Carter method for the United States, Japan, Canada, Sweden, and France systematically underestimated future life expectancy.

There have been other time-series methods to estimate and project age-specific mortality rates. The Brass relational method fits a two-parameter model where the age-specific mortality rates are assumed to be a linear function of a user-chosen model life table on a logit scale (Brass, 1971). The Heligman-Pollard model is an eight-parameter model with three parts describing mortality at different age ranges, childhood, young adult, and late-life (Heligman & Pollard, 1980). Although both models have been effective in fitting mortality data (e.g., Keyfitz, 1991; Hartmann, 1987), difficulties may arise in projecting the parameters (Keyfitz, 1981). Like the Heligman-Pollard model, the Bongaarts shifting logistic model (Bongaarts, 2005) differentiates mortality at different ages by fitting a 3-parameter logistic model and fixing the slope parameter across time while allowing the other two parameters to vary with time. However, the shifting logistic model focuses on senescent mortality and is relevant only in context in which infant/child and adult mortality are already negligible. Other models have focused on senescent mortality in terms of biological and evolutionary phenomena which incorporate the idea of heterogeneity in the population and frailty (See for example Steinsaltz & Wachter, 2006 and Yashin et al., 2000 for a summary of such models).

But again, these approaches require age-specific data.

Instead of modeling age-specific mortality rates, Gage (1993) used a competing hazards model developed by Siler (1979), which was developed of animal mortality. The five-parameter model describes the hazard function as three components, immmature (or childhood), residual (or background), and sensecent mortalities. Bongaarts (2006) also suggests the decomposition of mortality into these three components with future projections focusing on the senescent mortality component. Unfortunately, to properly fit both approaches require cause of death data that even when available for countries which collect such data, Gage (1993) himself acknowledges have limitations due to data quality and cause-of-death classification.

In addition to time-series approaches, there are two other main approaches to developing predictive distributions of projections (Lee, 1998). As was previously discussed, expert-based probabilistic projections have been produced by Lutz and colleagues at the IIASA (Lutz et al., 1998, 2004, 2008). However, this method does not explicitly rely on the use of available data, instead relying on a collection of diverse experts and their ability to specify specific probabilistic bounds, which may or may not be accurate (Alho, 2005). The other alternative to time-series methods is ex-post analysis of previous projections (Keyfitz, 1981; Stoto, 1983; Smith & Sincich, 1990). In this method, previous forecast errors are used to create probabilistic errors on future projections.

Girosi & King (2008) recently proposed a Bayesian method which incorporates covariates in a linear regression model. However, their approach depends on additional data which may not be reliable or even available in many countries. Pedroza (2006) proposed a Bayesian approach to the Lee-Carter approach by accounting for the uncertainty in the age parameters as well as the time parameter usually forecasted. Czado et al. (2005) also present a Bayesian approach to the Poisson log-bilinear formulation of the Lee-Carter model. While the latter two approaches account for uncertainty in the Lee-Carter model, the generalization of these

methods to all countries are again hindered by the data availability of age-specific mortality rates.

7.1 Future Research

For this initial analysis, we restricted the countries to those without generalized HIV/AIDS epidemics. For a secondary analysis, we loosened the exclusion rule and fitted our model to all countries with a WPP 2000-2005 HIV/AIDS prevalence rate of less than 4% (n=179). The 10-year cross-validation of these countries indicate the model maintains its predictive ability with 80% prediction intervals accurately predicting life expectancy 84% of the time. This continued calibration does not decrease the precision of the model. The mean absolute error of 1.2 and average 80% PI half-length of 1.9 years, which are only slightly higher than the results (1.0 and 1.7, respectively) from the non-generalized epidemics analysis. Further research is needed to generalize our model for countries with a generalized HIV/AIDS epidemic while properly accounting for the uncertainty in AIDS mortality, but our secondary analysis indicates generalization would be possible.

Further research is needed to apply this model to life expectancy at birth among women while ensuring trajectories by gender do not diverge or cross. This could potentially be done by modeling the two genders independently, as is recommended in the Lee-Carter method (Lee & Carter, 1992), and introducing a new parameter ensuring stochastic trajectories do not cross or diverge. The model could also be made more complex by allowing the double-logistic parameters to be correlated across genders.

In developed nations, age-specific mortality rates can be accurately estimated by vital registration and official censuses. However, this is not the case in a large percentage of the world where estimates are based on infrequent census and demographic surveys. The Millennium Development Goals to reduce child mortality has improved data collection and

estimation of infant and child mortality. But, the collection of adult mortality data is still sparse. Because of the inequities in data, it is important for future mortality projections to incorporate all sources of uncertainty using reproducible methodology for all nations.

To prevent overly optimistic long-range projections of life expectancy, we could assume that once a trajectory reaches a maximum, there will not be a mean increase in life expectancy. Instead, the trajectory would continue as a random walk without drift. This approach is a stochastic analog to the common practice of setting a deterministic maximum to life expectancy. A potential limit could be found in a recent Japanese life table (Wilmoth & Shkolnikov, 2009), where there was a 5% survival rate to the age of 96. Figure 8 shows life expectancy projections for Japan with such a maximum set. Although the median projected life expectancy does not surpass 96 in 90 years, the top half of the distribution in 2100 allows for possible life expectancies above 96.

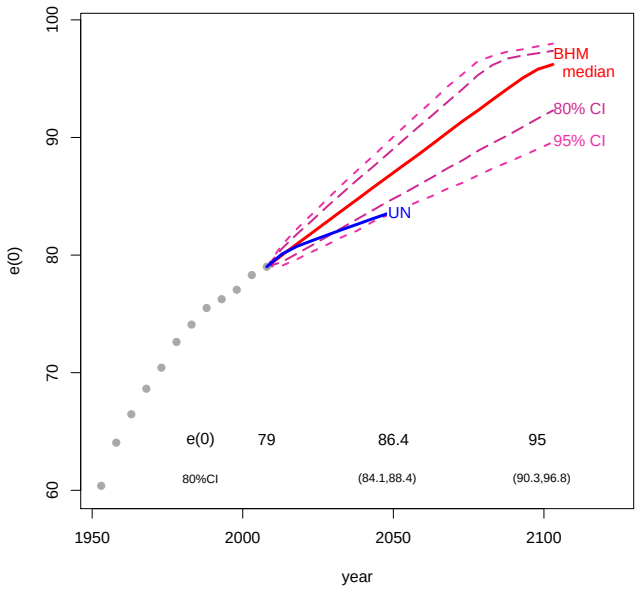


Figure 8: Life expectancy projections for Japan with the average life expectancy set to be no more than 96.

Assuming a significant declining rate of life expectancy gains other than what has been observed is not without potential dangers. In a recent study, Olshansky et al. (2009) found that current official forecasts by the U.S. Social Security Administration and Census Bureaus for male life expectancy in 2050 may be *underestimated* by 3.1 years (and 7.9 years for female life expectancy). They estimate this discrepancy could cost as much as \$3 to \$8 *trillion* more than currently projected for Medicare and Social Security. Using the discussion of potential medical breakthroughs that would slow aging to motivate potential shifts in age-specific mortality rates, Olshansky et al. (2009) produced numerous scenario-based projections of life expectancy for varying levels of TFR. The current official government agency forecasts for male life expectancy in 2050 are 80−81 years, whereas, Olshansky et al. (2009) projected 83−86 years under their two main mortality scenarios, with the range of 78-113 years in very extreme mortality scenarios. Our BHM projections are remarkably in line with Olshansky et al. (2009) with our median projection for 2050 at 84.5 years and an 80%CI of 82.2−86.4. The potentially large monetary repercussion of life expectancy projections clearly motivate the need for fully probabilistic projections life expectancy, as well as other demographic indicators.

Dowd et al. (2010) note that three types of uncertainty to be aware of when forecasting life expectancy, the model, parameter, and forecast uncertainties. With our approach, we updated the current UN model to be more flexible for fitting country-specific trends while accounting for the uncertainty in the parameter estimates in the double-logistic modeling of gains in life expectancy. We are then able to produce fully probabilistic projections of life expectancy that are demonstrably well-calibrated, without sacrificing the precision of the forecast intervals.

References

Alho, J. (2008). Aggregation across countries in stochastic population forecasts. *International Journal of Forecasting*, *24*(3), 343–353.

Alho, J. M. (2005). *Statistical demography and forecasting*. New York: Springer.

Bell, W. (1997). Comparing and assessing time series methods for forecasting age-specific fertility and mortality rates. *Journal of Official Statistics*, *13*, 279–202.

Bongaarts, J. (2005). Long-range trends in adult mortality: Models and projection methods. *Demography*, *42*(1), 23–49.

Bongaarts, J. (2006). How long will we live? *Population and Development Review*, *32*(4), 605–628.

Booth, H. (2006). Demographic forecasting: 1980 to 2005 in review. *International Journal of Forecasting*, *22*(3), 547–581.

Booth, H., Tickle, L., & Smith, L. (2005). Evaluation of the variants of the Lee-Carter method of forecasting mortality: A multicountry comparison. *New Zealand Population Review, Special Issue on Stochastic Population Projections*, *31*(13-37).

Brass, W. (1971). *Biological aspects of demography*, chapter On the scale of mortality, (pp. 69–110). London: Taylor and Francis.

Brouhns, N., Denuit, M., & Vermunt, J. (2002). A poisson log-bilinear regression approach to the construction of projected life tables. *Insurance: Mathematics and Economics*, *31*, 373–397.

Czado, C., Delwarde, A., & Denuit, M. (2005). Bayesian poisson log-bilinear mortality projections. *Insurance: Mathematics and Economics*, *36*(3), 260 – 284.

de Jong, P. & Tickle, L. (2006). Extending Lee-Carter mortality forecasting. *Mathematical Population Studies*, *13*, 1–18.

Dowd, K., Blake, D., & Cairns, A. J. G. (2010). Facing up to the uncertain life expectancy:the longevity fan charts. *Demography*, *47*(1), 67–78.

Fries, J. F. (1980). Aging, natural death, and the compression of morbidity. *New England Journal of Medicine*, *202*(3), 130–135.

Gage, T. B. (1993). The decline of mortality in england and wales 1861 to 1964: Decomposition by cause of death and component of mortality. *Population Studies*, *47*(1), 47–66.

Gelman, A. & Rubin, D. (1992). Inference from iterative simulation using multiple sequences (with discussion). *Statistical Science*, *7*, 457–511.

Girosi, F. & King, G. (2008). *Demographic forecasting*. Princeton, NJ: Princeton University Press.

Gneiting, T., Balabdaoui, F., & Raftery, A. (2007). Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society, Series B*, *69*, 243–268.

Hartmann, M. (1987). Past and recent attempts to model mortality at all ages. *Journal of Official Statistics*, *3*, 19–36.

Heligman, L. & Pollard, J. H. (1980). The age pattern of mortality. *Journal of the Institute of Actuaries*, *107*, 49–80.

Hyndman, R. J. & Ullah, S. (2007). Robust forecasting of mortality and fertility rates: A functional data approach. *Computational Statistics and Data Analysis*, *51*(10), 4942–4956.

Ishii, F. (2008). Mortaliy projection model for Japan with age-shifting structure. Paper presented at the annual meeting of The Population Association of America, April 17-19, New Orleans, Louisiana.

Kaneko, R., Ishikawa, A., Ishii, F., Sasai, T., Iwasawa, M., Mita, F., & Moriizumi, R. (2008). Population projections for Japan: 2006-2055 outline of results, methods, and assumptions. *The Japanese Journal of Population*, *6*(1), 76–114.

Keyfitz, N. (1981). The limits of population forecasting. *Population and Development Review*, *7*, 579–593.

Keyfitz, N. (1991). Experiments in the projection of mortality. *Canadian Studies in Population*, *18*, 1–17.

Koissi, M., Shapiro, A., & Hgns, G. (2006). Evaluating and extending the lee-carter model for mortality forecasting: Bootstrap confidence interval. *Insurance: Mathematics and Economics*, *38*, 1–20.

Lee, R. & Miller, T. (2001). Evaluating the performance of the Lee-Carter method for forecasting mortality. *Demography*, *38*(4), 537–549.

Lee, R. D. (1998). Probabilistic approaches to population forecasting. *Population and Development Review*, *24*, 156–190.

Lee, R. D. & Carter, L. R. (1992). Modeling and forecasting U. S. mortality. *Journal of the American Statistical Association*, *87*(419), 659–671.

Li, N. & Lee, R. (2005). Coherent mortality forecasts for a group of populations: An extension of the Lee-Carter method. *Demography*, *42*(3), 575–594.

Li, N., Lee, R., & Tuljapurkar, S. (2004). Using the Lee-Carter method to forecast mortality for populations with limited data. *International Statistical Review*, *72*, 19–36.

Lunn, D., Thomas, A., Best, N., & Spiegelhalter, D. (2000). WinBUGS v1.4.3 – a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing*, *10*, 325–337.

Lutz, W., Sanderson, W. C., & Scherbov, S. (1998). Expert-based probabilistic population projections. *Population and Development Review*, *24*, 139–155.

Lutz, W., Sanderson, W. C., & Scherbov, S. (2004). *The end of world population growth in the 21st century: New challenges for human capital formation and sustainable development*. Sterling, VA: Earthscan.

Lutz, W., Sanderson, W. C., & Scherbov, S. (2008). IIASAs 2007 probabilistic world population projections, IIASA world population program online data base of results. Retrieved November 6, 2009 (http://www.iiasa.ac.at/Research/POP/proj07/index.html?sb=5).

Meyer, P. (1994). Bi-logistic growth. *Technological Forecasting and Social Change*, *47*, 89–102.

Neal, R. M. (2003). Slice sampling. *The Annals of Statistics*, *31*(3), 705–741.

Oeppen, J. & Vaupel, J. (2002). Broken limits to life expectancy. *Science*, *296*, 1029–1031.

Olshansky, S.J., B. C. & Dsesquelles, A. (2001). Prospects for human longevity. *Science*, *291*, 1491–1492.

Olshansky, S., Carnes, B., & Cassel, C. (1990). In search of methuselah: Estimating the upper limits to human longevity. *Science*, *250*, 634–640.

Olshansky, S., Passaro, D., Hershow, R., Layden, J., Carnes, B., Brody, J., Hayflick, L., Butler, R., Allison, D., & Ludwig, D. (2005). A potential decline in life expectancy in the united states in the 21st century. *New England Journal of Medicine*, *352*, 1103–1110.

Olshansky, S. J., Carnes, B., & Brody, J. (2002). A biodemographic interpretation of the life span. *Population and Development Review*, *28*.

Olshansky, S. J., Goldman, D. P., Zheng, Y., & Rowe, J. W. (2009). Aging in America in the twenty-first century: Demographic forecasts from the MacArthur Foundation Research Network on an Aging Society. *The Milbank Quarterly*, *87*(4), 842862.

Pedroza, C. (2006). A bayesian forecasting model: predicting u.s. male mortality. *Biostatistics*, *7*(4), 530–550.

R Development Core Team (2009). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.

Raftery, A. & Lewis, S. (1992). One long run with diagnostics: Implementation strategies for Markov chain Monte Carlo. *Statistical Science*, *7*, 493–497.

Renshaw, A. & Haberman, S. (2006). A cohort-based extension to the Lee-Carter model for mortality reduction factors. *Insurance: Mathematics and Economics*, *38*, 556–570.

Siler, W. (1979). A competing-risk model for animal mortality. *Ecology*, *60*(4), 750–757.

Smith, S. K. & Sincich, T. (1990). On the relationship between length of base period and population forecast errors. *Journal of the American Statistical Association*, *85*, 367–375.

Steinsaltz, D. R. & Wachter, K. W. (2006). Understanding mortality rate deceleration and heterogeneity. *Mathematical Population Studies*, *13*, 19–37.

Stoto, M. A. (1983). The accuracy of population projections. *Journal of the American Statistical Association*, *78*, 13–20.

Sturtz, S., Ligges, U., & Gelman, A. (2005). R2WinBUGS: a package for running WinBUGS from R. *Journal of Statistical Software*.

Tuljapurkar, S. (2005). Future mortality: A bumpy road to shangri-la? *Science of Aging Knowledge Environment*, *14*, pe9.

Tuljapurkar, S., Li, N., & Boe, C. (2000). A universal pattern of mortality decline in the g7 countries. *Nature*, *405*, 789–792.

United Nations (2006). *World Mortality Report 2005*. New York, NY: United Nations.

United Nations (2009). *World population prospects: The 2008 Revision*. New York, NY: United Nations.

Vallin, J. & Meslé, F. (2009). The segmented trend line of highest life expectancies. *Population and Development Review*, *35*(1), 159–187.

Vaupel, J., Carey, J., Christensen, K., Johnson, T., Yashin, A., Holm, V., Iachine, I., Kannisto, V., Khazaeli, A., Liedo, P., Longo, V., Zeng, Y., Manton, K., & Curtsinger, J. (1998). Biodemographic trajectories of longevity. *Science*, *280*, 855–860.

White, K. M. (2002). Longevity advances in high-income countries, 1955-96. *Population and Development Review*, *28*(1), 59–76.

Wilmoth, J. R. & Shkolnikov, V. (2009). The human mortality database:Japan male life table. `http://www.mortality.org`.

Yashin, A. I., Iachine, I. A., & Begun, A. S. (2000). Mortality modeling: a review. *Mathematical Population Studies*, *8*(4), 305–332.