Bayesian Probabilistic Projections of Mortality DRAFT

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

May 4, 2009

Contents

*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 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.

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 estimates 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 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.

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.

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) 2006 Revision from 1950 through 2005 (Nations, 2007). 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 article, 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 then discuss various metrics for assessing the predictive power of our model via cross-validation. In the next section, we present results of the cross-validation with an illustrative example from Uzbekistan. Comparisons are then made with the recently updated regional projections for South Asia by IIASA. Lastly, we present projections for Japan, the leading country in overall life expectancy.

2 Data

Because life expectancy at birth is a summary indicator for all age-specific mortality rates, 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 2007 UN World Mortality Report, since 1990, only 56% of 195 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

[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

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 2006 Revision (WPP) produced by the United Nations Population Division (Nations, 2007). WPP estimates are an expert culmination of the disparate data and methodological machinery available.

The HIV/AIDS epidemic major source of mortality in the last 20 years. As such, we exclude countries with a generalized HIV/AIDS epidemic [2] as was denoted on UNAIDS fact sheets. A country is defined as 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%. Table 1 lists the excluded countries.

3 Methodology

3.1 Model

Currently, the UN estimates 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 157 countries from 1950 to 2000 are plotted in Figure

[2] Classification of countries with a generalized HIV/AIDS epidemic was based on HIV/AIDS fact sheets published jointly by WHO, UNAID and UNICEF (2008).

Table 1: Countries with a generalized HIV/AIDS epidemic as denoted by UNAIDS fact sheets and 2000-2005 prevalence rate (WPP 2006).

Country	Prev Rate	Country	Prev Rate
Niger	1.1	Guinea-Bissau	3.8
Guinea	1.5	Nigeria	3.9
Sierra Leone	1.6	Congo	5.3
Mali	1.7	Cameroon	5.4
Benin	1.8	Kenya	6
Ethiopia	1.9	Uganda	6.4
Burkina Faso	2	United Rep. of Tanzania	6.4
Ghana	2.3	Cte d'Ivoire	7
Eritrea	2.4	Gabon	7.8
Gambia	2.4	Central African Rep.	10.7
Rwanda	3	Malawi	14
Djibouti	3.1	Mozambique	16.1
Equatorial Guinea	3.2	Zambia	16.9
Togo	3.2	South Africa	18.7
Dem. Rep. of the Congo	3.2	Namibia	19.3
Burundi	3.2	Zimbabwe	20.1
Liberia	3.4	Lesotho	23.2
Chad	3.5	Botswana	26.5
Angola	3.7	Swaziland	34.2

3.1. This figure highlights the non-constant rate of change in life expectancy. To capture this, the UN has developed a model that represents the decline in mortality by fitting a double-logistic function of current life expectancy.

The double-logistic function has six parameters, as is illustrated in Figure 3.1, 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 by assessing the pace of mortality decline in the recent to medium-term past (Nations, 2009). In the UN approach, a constant rate is set once the gains in life expectancy have reached a preset low level. Al-

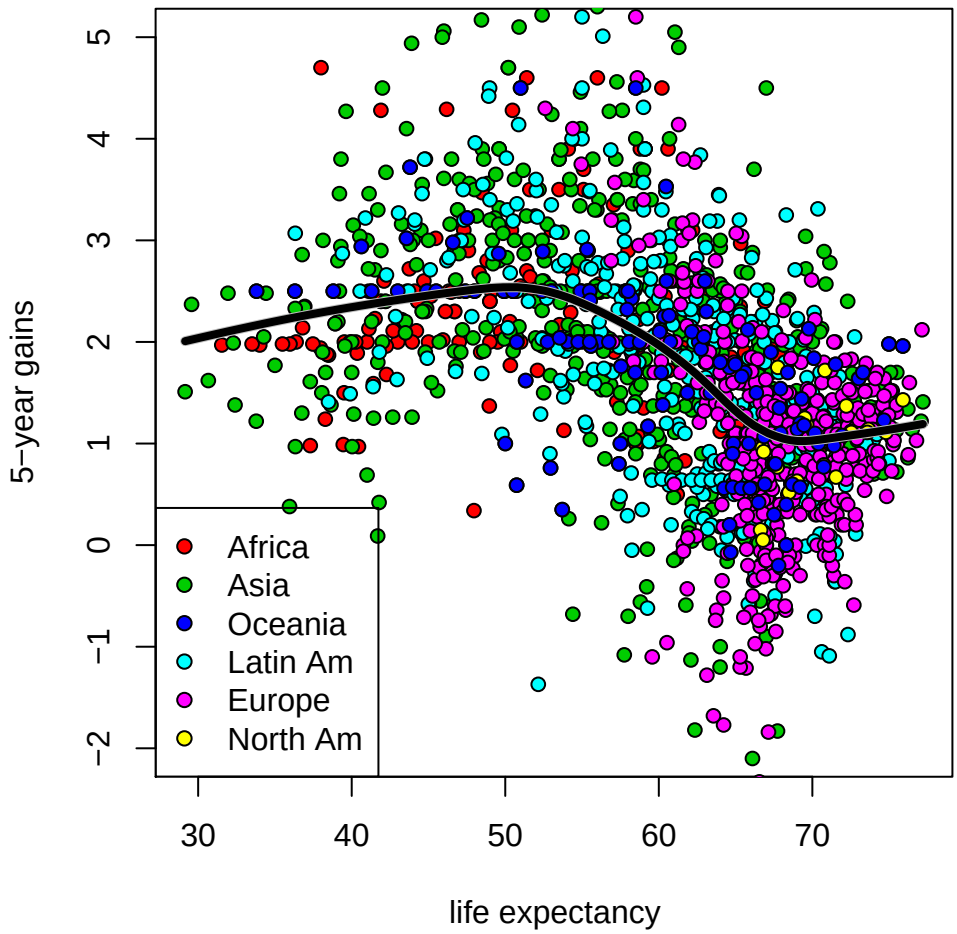

Figure 1: Each point represents the observed five-year gain in life expectancy within a country. The green line is a locally-weighted polynomial regression (lowess) of the observations. UN estimates for 157 countries from 1950 to 2000 are included in this figure (n=1256). (Note, 24 observations (1.9%) are outside the range of the plot and not depicted, but where include in the local regression.)

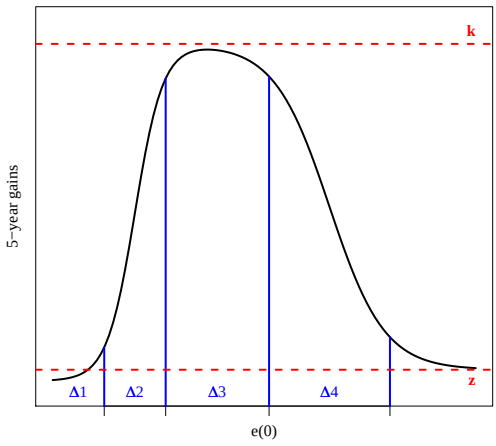

Figure 2: Illustration of the double-logistic parameters (blue).

though 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. Demographic transition theory suggests slow increase 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

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

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

$$
\begin{aligned}
g(l_{c,t}|\boldsymbol{\theta}^{(UN)}) &= \frac{\theta_5^{(UN)}}{1+exp(-\frac{log(9^2)}{\theta_2^{(UN)}}(l_{ct}-m_1^{(UN)}))} + \frac{\theta_6^{(UN)}-\theta_5^{(UN)}}{1+exp(-\frac{log(9^2)}{\theta_4^{(UN)}}(l_{ct}-m_2^{(UN)}))} \\
\boldsymbol{\theta}^{(UN)} &\in (\boldsymbol{\theta}^{\text{Very Slow}}, \boldsymbol{\theta}^{\text{Slow}}, \boldsymbol{\theta}^{\text{Medium}}, \boldsymbol{\theta}^{\text{Fast}}, \boldsymbol{\theta}^{\text{Very Fast}}) \\
\boldsymbol{\theta}^{(UN)} &= (\theta_1^{(UN)}, \theta_2^{(UN)}, \theta_3^{(UN)}, \theta_4^{(UN)}, \theta_5^{(UN)}, \theta_6^{(UN)})
\end{aligned}
$$

where,

$$
\begin{aligned}
m_1^{(UN)} &= \theta_1^{(UN)} + 0.5\theta_2^{(UN)} \\
m_2^{(UN)} &= \theta_1^{(UN)} + \theta_2^{(UN)} + \theta_3^{(UN)} + 0.5\theta_4^{(UN)}.
\end{aligned}
$$

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 ($\delta_{c,t+1}$):

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

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. A Bayesian-based approach allows the estimation of

country-specific probabilistic distributions of gains in life expectancy. Our model is given by

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

$$\begin{aligned} g(l_{c,t}|\boldsymbol{\theta}^{(c)}) &= \text{Double-Logistic function with parameters } \boldsymbol{\theta}^{c} \\ \boldsymbol{\theta}^{(c)}|\Sigma &\sim \text{Normal}_{Trunc}(\boldsymbol{\theta}, \Sigma) \\ \Sigma &= \text{diag}(\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2, \sigma_5^2, \sigma_6^2). \end{aligned}$$

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. By assuming the sixth parameter, z, is non-negative, we are assuming that on average, life expectancy will continue to increase. This is consistent with the findings of Oeppen & Vaupel (2002) where they determined there was no indication that life expectancy would stop increasing.

Figure 3.1 shows the standard deviation of the residuals when we assume homoskedasticity around the double-logistic function. As is evident from the locally smoothed, Loess, regression line of these statistics, the observations 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 (κ) to the loess fit of life expectancy versus the standard deviation of residuals of all countries ($r(l_{c,t})$). We model the

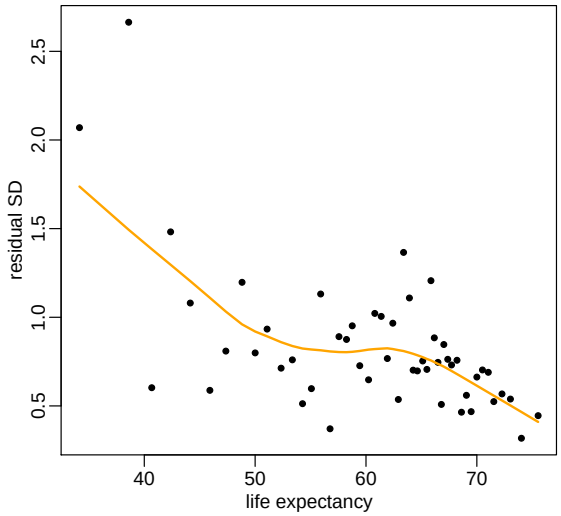

Figure 3: The standard deviation of residuals across life expectancy with locally weighted regression fit.

stochastic error term around the double-logistic function as

$$\delta_{ct} \sim_{iid} N(0, (\kappa \times r(l_{c,t-1}))^2) \tag{4}$$

3.2 Parameter estimation

To carry out a Bayesian analysis, we specify the probabilistic distributions for all parameters in the model as follows:

$$\begin{aligned}
\sigma_\delta^2 &\sim \text{Inverse-Gamma}(\frac{\nu_\delta}{2}, \frac{\nu_\delta \omega_\delta^2}{2}) \\
\boldsymbol{\theta}|\Sigma &\sim \text{Normal}_{Trunc}(\boldsymbol{\theta}_0, \alpha\Sigma) \\
\sigma_{\theta,i}^2 &\sim \text{Inverse-Gamma}(\frac{\nu_{\theta,i}}{2}, \frac{\nu_{\theta,i}\omega_{\theta,i}^2}{2}) \\
\kappa &\sim \text{Uniform}(a, b).
\end{aligned}$$

Prior specification of the world-level parameters were assumed to be sampled from a truncated normal density with variance parameters independently sampled from an inverse-gamma distribution. Empirical Bayes method of hyper-parameter specification with priors having the weight of one observation within one country ($nu_\delta, \alpha, \nu_{\theta,i} = 1$). For the stochastic proportionality constant, κ, a diffuse prior was specified uniformally from zero to ten.

Models were run using the package R2WinBUGS (2005) in R 2.7 (2009) to access WinBugs 1.4 (2000). The posterior distribution of the parameter was sampled via slice sampling Markov Chain Monte Carlo (MCMC) methods. (See REF for information about MCMC methods.) For each run of the model, three chains of length 10,000 were run with a burn-in of 1,000 with a thinning factor of 2 resulting in three simultaneous chains of length 4,500. Standard diagnostic, which are available in the R package coda (Plummer et al.), all 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 & Raftery (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=1,413) and forecasting life expectancy for males from 1995 to 2005, resulting in 314 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 2. Overall, our model was well calibrated with our 95% prediction intervals capturing the actual observations 95% of the time. The nominal coverage of our 80% prediction intervals was 84%. The mean standardized absolute error (SAPE) was 0.96, which is quite close to the theoretical mean of 1. The mean absolute error (MAE) of our median predictions was 1.00 year. That is, over 10 years of predictions, our "best guess", on average, was within 1 year of the actual observation

Table 2: Summary measures for 10 year out-of-sample cross-validation.

Summary statistics		
Root mean square error (MSE)		1.72
Std absolute prediction error (SAPE)		0.96
Mean absolute error (MAE)		1.00
Prediction Intervals		
Nominal	Actual	Mean length
95%	95%	±2.6
90%	92%	±2.2
80%	84%	±1.7

Gneiting & Raftery (2007), propose 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, 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. Figure 4 contains boxplots of half the lengths of the 80% prediction interval for the two projected quinquennial periods overall and broken down by (UN-defined) region. Excluding outliers, for the 1995-2000 time period, the prediction interval half-lengths range from 0.8 to 1.9 years with an average length of 1.4 years. For the next quinquennial period, 2000-2005, the interval half-lengths increase to a range of 1.1 to 2.7 years with an average half-length of 2 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 and average interval half-length of 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. Looking across regions, we can see that as life expectancy increases, the prediction intervals are more narrow, indicating there is less variability in life expectancy as it increases.

4.1 Country-specific projections for Uzbekistan (Out of sample)

Figure 4.1 is a plot of estimated and projected life expectancy for Uzbekistan. The actual observed UN WPP 2006 time-series is indicated by solid black circles with projections in dark blue. The UN WPP 1996 projections for 1995-2050 are represented by the solid light blue line and ours are in red with the 80%PI represented with dashed lines. This allows us to rewind time to see how accurate our projections and the UN WPP 1996 projections would have been.

Male life expectancy in Uzbekistan was increasing from 52.5 years in 1950 to 64 years in 1990. However, in the next quinquennial, 1990-1995, male life expectancy in Uzbekistan *decreased* by 1 year to 63 years. Both WPP 1996 and our median projections predict a

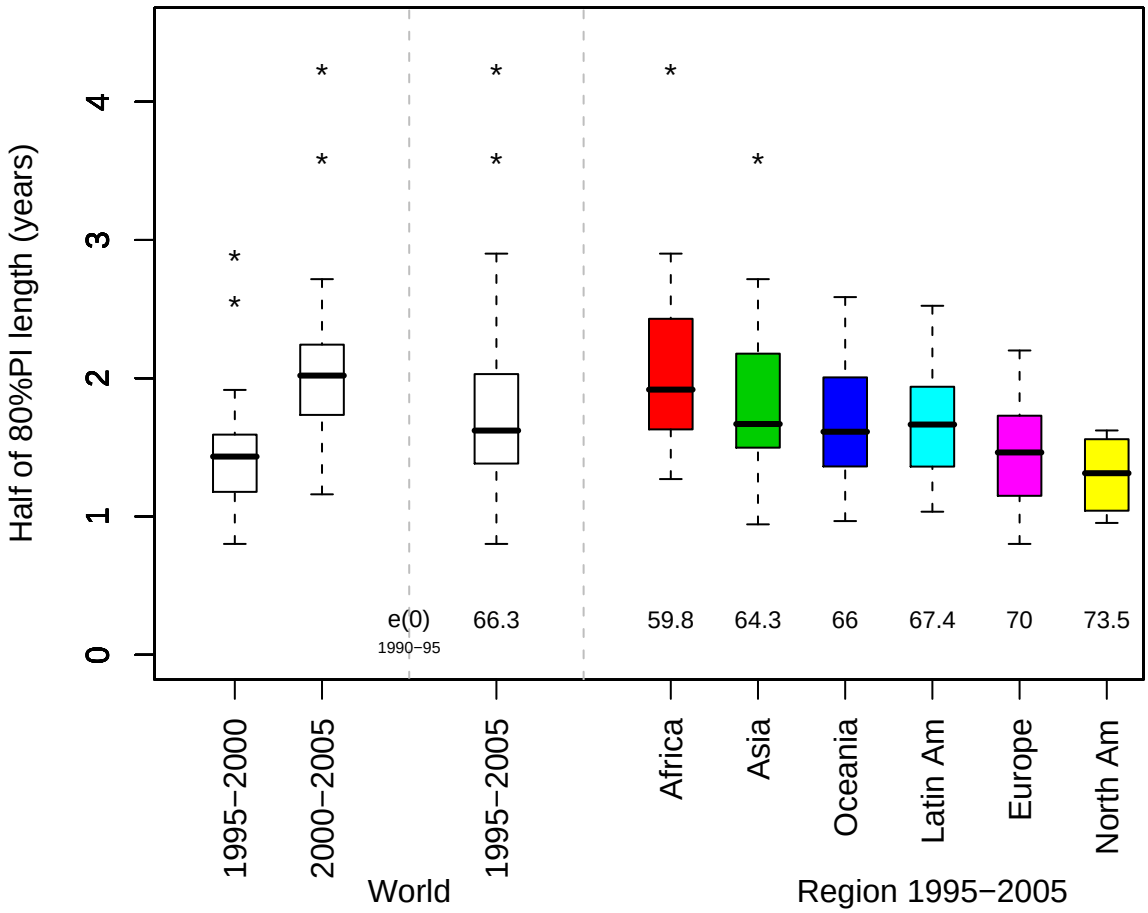

Figure 4: Boxplots of half the 80% prediction interval lengths from the out-of-sample projections of life expectancy from 1995-2005.

continuous increase in life expectancy, however, ours increase at a more conservative rate. As is evident in Figure 4.1, our 80% prediction intervals capture the "true" (Nations, 2007) 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 66.1 years, which is about half a year higher than WPP 1996 (Nations, 2007) estimate of 64.5. Yet, the lower bound of our 80% prediction interval, 61.9 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 25 years, which highlights the uncertainty of the previous time-series.

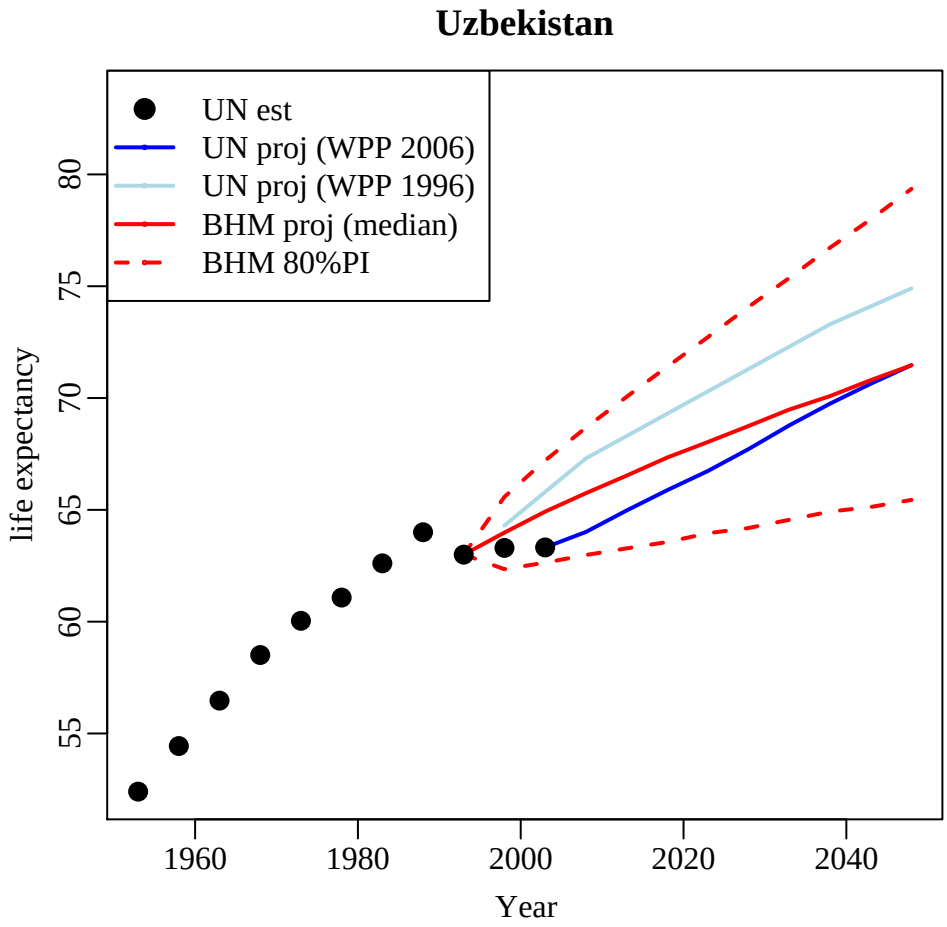

Figure 5: Life expectancy for Uzbekistan with our out-of-sample projections (and 80%PI) starting in 1995 (red), UN WPP 1996 (light blue) and UN WPP 2006 projections (dark blue).

4.2 Regional projections

As previously mentioned, researchers at the International Institute for Applied Systems Analysis (IIASA) (Lutz et al. 2004; Lutz, et al, 1996 1997, 2001) produced regional probabilistic prediction intervals for life expectancy using Delphi-type methods. A range of experts were asked to give 90% prediction intervals for future life expectancy in each of 13 specified regions. 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, we aggregated UN estimates and projections and our projections to be proportional to the regional male populations in 2003. Note, we assume the life expectancy projections are independent. [3] Positive intra-regional correlations may increase the uncertainty in our regional projections.

As defined by IIASA, the countries, male population and weight value for South Asia are in Table 3. Figure 4.2, includes projections of life expectancy for the South Asian region. The 2007 IIASA projections available on their website (Lutz, Sanderson & Scherbov, Lutz et al.) are depicted. The decennial intervals lengths, in Figure 4.2, indicate the IIASA intervals are 32.5% more narrow than our projections in 2008. However, our projections begin 4 years earlier in 2003, the mid-year of the 2000-2005 quinquennial. Regardless of this discrepancy, for subsequent projections, our intervals are more sharp, ranging from 12-74% more narrow than those of IIASA.

[3] See Alho (2008) for a discussion on aggregation across countries in the European Union.

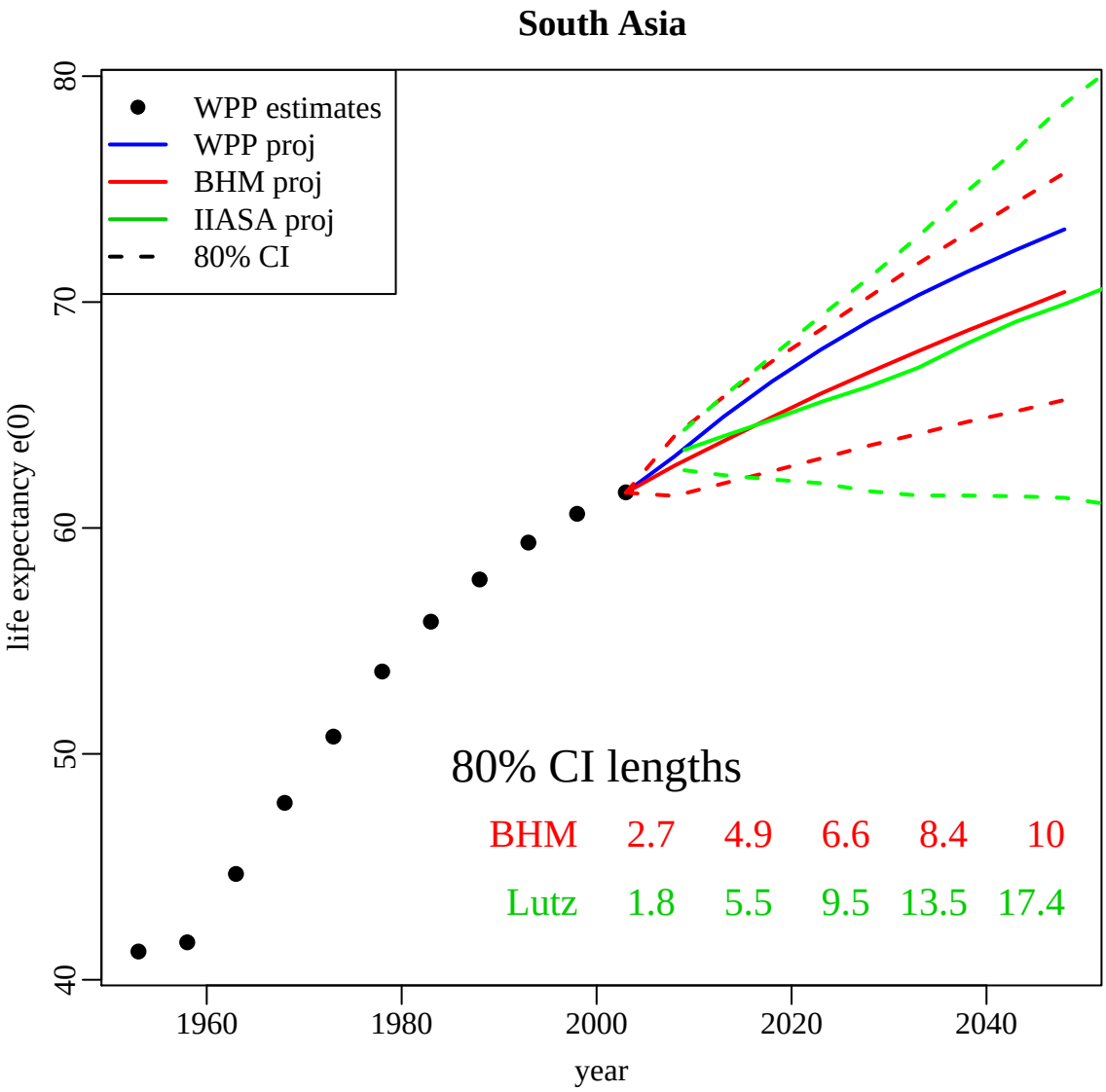

Figure 6: Life expectancy projections for South Asia (IIASA-defined) for our model (red), IIASA (green) and WPP (blue).

Table 3: Male Population and relative proportion of countries within South Asian (IIASA-defined).

South Asia		
Country	Male Population (1,000s)	Pct of Region Pop (%)
India	522.2	75.1
Pakistan	72.8	10.48
Bangladesh	68.8	9.89
Nepal	11.6	1.67
Afghanistan	10.2	1.47
Sri Lanka	9.3	1.33
Bhutan	0.3	0.04
Maldives	0.1	0.02

4.3 Projecting leading countries

One of the difficulties with projecting mortality is accurately projecting the country with the lowest mortality. Historically, many assumed there must be a "ceiling" to life expectancy for humans (REFS). However, past estimates of the "maximum life expectancy" have continually been surpassed (Oeppen and Vauppel, 2002, OTHER REFS). In fact, Oeppen & Vaupel (2002) presented strong evidence that the worlds 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 (that title currently belongs to Iceland), it is the country with the highest overall life expectancy and has been for many years. Figure 4.3 is a plot of male life expectancy in Japan. The green line is what the trajectory would be if male life expectancy in Japan increased at the "best practices" rate. Our median projection, indicated as a red solid line, are strikingly consistent with the "best practices" trajectory.

Recently, both the UN and the Japanese official projections made by the National Institute of Population and Social Security Research (IPSS) 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. IPSS and UN now use the shifting logistic model (Bongaarts, 2005) to account for continued increase in life expectancy in Japan. IPSS projections (low/medium/high rates of mortality decline variants) (Kaneko et al., 2008) are included in Figure 4.3 in light blue. The IPSS are more conservative than both our and the "best practices" projections, yet still within our prediction interval.

5 Discussion

Much research has been done on the forecasting of mortality (see Booth, 2006 and OTHERS 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; Perosi 2006) and extensions (e.g., Li & Lee, 2005; Li et al., 2004).

There have been other time-series methods to estimate and project age-specific mortality rates, including the Heligman-Pollard model (Heligman & Pollard, 1980) and Brass methods (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. 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. 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).

In addition to time-series approaches, there are two other main approaches to developing predictive distributions of projections (Lee, 1998; National Research Council 2000). As was

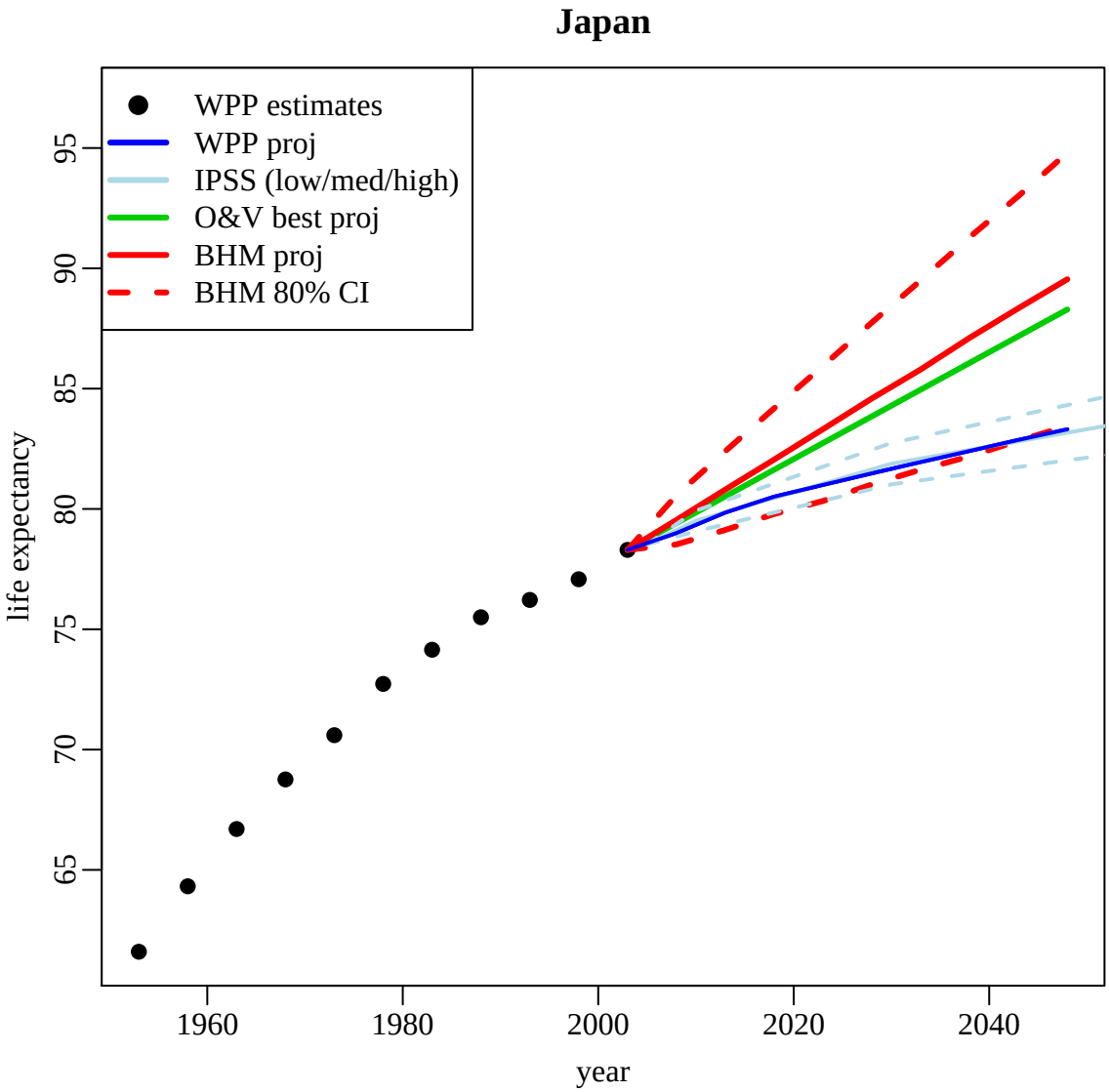

Figure 7: Life expectancy projections for Japan, which has the highest overall life expectancy. The UN WPP projections are in blue. National Institute of Population and Social Security Research (IPSS) variants in light blue. The green line indicates a constant increase of 0.222 per year, as found by Oeppen and Vauppel (O & V) (2002) among the best practices country. Our probabilistic projections and 80% CI are indicated in red.

previously discussed, expert-based probabilistic projections have been produced by Lutz and colleagues at the IIASA (1998). 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.

The use of Bayesian framework . 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. Perosi (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.

5.1 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=178). The 10-year cross-validation of these countries indicate the model maintains its predictive ability with 80% prediction intervals accurately predicting life expectancy 82% 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 1.8 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.

The model we presented here assumed that the random perturbations 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. Our approach can be adapted to modeling the perturbations, $\delta_{c,t}$, as samples from a multivariate distribution where correlations within a region are also estimated. However, this approach would require the *a priori* definition of each region and a proposed covariance structure (e.g., spatial correlations, correlations based on population size). Our model can also be further adapted to be more complex ARIMA model.

Further research is needed to apply this model to life expectancy at birth among women while ensuring trajectories be 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 accurately be 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.

6 List of Notation

Table 4: Model Notation

c	country, $c \in 1, 2, \ldots, 157$
$g(l_{ct})$	double-logistic function modeling the expected 5-year gains in life expectancy for country c at time period t
$l_{c,t}$	life expectancy at birth in country c during time period t
$r(l_{ct})$	function accounting for heteroscedasticity in perturbations
t	quinquennial time period
δ_{ct}	random perturbations
$\boldsymbol{\theta}$	vector of world-level double-logistic parameters $(\theta_{1i}, \theta_{2i}, \theta_{3i}, \theta_{4i}, \theta_{5i}, \theta_{6i})$
$\boldsymbol{\theta}_c$	vector of country-specific double-logistic parameters
κ	proportionality constant of perturbations
Σ	Variance-Covariance matrix of the world-level double-logistic parameters

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.

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 leecarter 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.

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

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

Gneiting, T., B. 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.

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*, 579593.

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

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 leecarter 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. Iiasas 2007 probabilistic world population projections, iiasa world population program online data base of results.

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.

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

Nations, U. (2007). *World population prospects: The 2006 Revision*. New York, NY: United Nations.

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

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

Plummer, M., Best, N., Cowles, K., & Vines, K. coda: Output analysis and diagnostics for mcmc.

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.

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.

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.

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*.