Outline

- Motivation
- Overview of population projection methods
- UN approach for probabilistic projections
  - Probabilistic fertility projections
  - Probabilistic mortality projections
- UN probabilistic population projections
- Summary
- Software and reference
Outline

- Motivation
- Overview of population projection methods
- UN approach for probabilistic projections
  - Probabilistic fertility projections
  - Probabilistic mortality projections
- UN probabilistic population projections
- Summary
- Software and reference
Why population projections?

- To assess hypothetical population trends based on specific assumptions about future trends in fertility, mortality and migration

- To help understand the determinants of population change and inform policy discussions

- To provide a base for other projections essential for social and economic planning (labor, education, social security, agriculture, health, housing, urbanization, energy, transport, climate, environment, etc.)

- To produce current demographic estimates using latest available data on population size (by age and sex) and its components of change (fertility, mortality and migration)

- To identify realistic goals and targets for future development trends
Future is unknown, but we know some basic demographic trends

- Demographic processes are **long-term**
  - Lasting impact of past and current changes
  - Population momentum

- Components of population change
  - Fertility
  - Mortality
  - Migration

- **Demographic transition** as guiding principle
  - Countries move from high to low levels of mortality and fertility
  - Still in progress in many developing countries
UN population projections

- UN Population Division publishes estimates and projections, by age and sex, of population counts and vital rates for all countries, for 5-year intervals of age and time, from 1950 to 2100, every two years in World Population Prospects (WPP)
  - Used throughout UN system, especially as denominators
  - Key input for development planning, monitoring (e.g. MDGs) and modeling (e.g. climate)
  - UN has produced 23 sets of global population projections since 1951
  - Latest version: the 2012 Revision, published in May 2013

- Population can be projected far into the future using current population by age, and age-specific rates of fertility, mortality, and migration
  - Governments often project over shorter intervals: 2060 (EU, USA, Japan), 2046 (Ireland)
  - UN projects to 2100 due to demand for long-term trends
Uncertainty

• Need some means of reflecting the uncertainty of population projections

• Different methods of depicting and/or measuring uncertainty
  • Describe a range of scenarios based on specific assumptions
  • Choose a central scenario and model the uncertainty around that scenario
  • Draw on the variability of expert predictions

• Major challenges in transmitting the meaning of uncertainty, especially to lay audiences
Outline

- Motivation

- **Overview of population projection methods**

- UN approach for probabilistic projections
  - Probabilistic fertility projections
  - Probabilistic mortality projections

- UN probabilistic population projections

- Summary

- Software and reference
Cohort component method

The method starts from current population estimates and projects population forward:

- **Demographic balancing equation:**
  \[ \text{Pop}(t+1) = \text{Pop}(t) + \text{Births}(t) - \text{Deaths}(t) + \text{Immig}(t) - \text{Emigr}(t) \]

- **Age-structured version:**
  \[ \text{Pop}(x+1,t+1) = \text{Pop}(x,t) \ast \text{Survival}(x,t) + \text{Net migr}(x,t) \]
  \[ \text{Pop}(0,t+1) = \sum \text{Women}(x,t) \ast \text{Fertility}(x,t) \]

- The cohort component method is based on **age-structured populations and components of change** (i.e., births, deaths, migration)
Cohort-component projections: step by step

Cohort-component projections: 5 years and over

- Deaths (+ I - E)
- Deaths (+ I - E)
- Pop(0-4)
- Pop(5-9)
- Pop(10-14)

Example with x=0
Variants and scenarios

- Uncertainty of future outcomes can be illustrated using variants and scenarios

- Variants describe a range of assumptions for a particular component of change (e.g. fertility), illustrating the sensitivity of outcomes to changes in assumptions

- Scenarios describe a series of hypothetical (often simplified) future trajectories, illustrating core concepts such as population momentum
### UN deterministic projection scenarios

8 scenarios included in the 2012 Revision of the UN World Population Prospects

<table>
<thead>
<tr>
<th>#</th>
<th>UN projection scenarios</th>
<th>Fertility variant</th>
<th>Mortality variant</th>
<th>International Migration variant</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Low fertility</td>
<td>Low</td>
<td>Normal</td>
<td>Normal</td>
</tr>
<tr>
<td></td>
<td>(= medium - 0.5 child)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td><strong>Medium fertility</strong></td>
<td>Medium</td>
<td>Normal</td>
<td>Normal</td>
</tr>
<tr>
<td>3</td>
<td>High fertility</td>
<td>High</td>
<td>Normal</td>
<td>Normal</td>
</tr>
<tr>
<td></td>
<td>(= medium + 0.5 child)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Constant-fertility</td>
<td>Constant as of 2005-2010</td>
<td>Normal</td>
<td>Normal</td>
</tr>
<tr>
<td>5</td>
<td>Instant-replacement-fertility</td>
<td>Instant-replacement as of 2010-2015</td>
<td>Normal</td>
<td>Normal</td>
</tr>
<tr>
<td>6</td>
<td>Constant-mortality</td>
<td>Medium</td>
<td>Constant as of 2005-2010</td>
<td>Normal</td>
</tr>
<tr>
<td>7</td>
<td>No change</td>
<td>Constant as of 2005-2010</td>
<td>Constant as of 2005-2010</td>
<td>Normal</td>
</tr>
<tr>
<td>8</td>
<td>Zero-migration</td>
<td>Medium</td>
<td>Normal</td>
<td>Zero as of 2010-2015</td>
</tr>
</tbody>
</table>
UN deterministic scenarios, total population: World 2010-2100
Uncertainty in demographic projections

Nathan Keyfitz (1981):

“Demographers can no more be held responsible for inaccuracy in forecasting population 20 years ahead, than geologists, meteorologists, or economists when they fail to announce earthquakes, cold winters, or depressions 20 years ahead. What we can be held responsible for is warning one another and our public what the error of our estimates is likely to be.”
Three approaches to probabilistic projections

- **Ex-post analysis** based on the errors in past forecasts (Keyfitz 1981; Stoto 1983; Alho 2006; Alders 2007; Alho 2008)

- **Time series methods** use past time series of forecast inputs, such as fertility and mortality, to estimate a statistical time series model, which is then used to simulate a large number of stochastic possible future demographic pathways. Simulated trajectories of forecast inputs are combined via a cohort component projection model to produce predictive distributions of forecast outputs (Lee 1994; Tuljapurkar 1999)

- **Expert-based approaches** rely on experts to provide distributions for each forecast input. These are then used to construct predictive distributions of forecast outputs using a stochastic method similar to the time series method (National Research Council 2000; Booth 2006; Pflaumer 1988; Lutz 1996, 1998, 2004)
Some limitations of expert opinion

- Difficulty of identifying a pool of experts qualified to provide quantitative predictions
- Difficulty of understanding and measuring potential biases in this pool
- Experts tend to overemphasize current situation and miss historical shifts
  - Baby Boom
  - HIV/AIDS
  - Fall of Soviet Union
  - Below-replacement fertility
Projections for England by ONS

Figure 2. Past Projections of Total Fertility Rates and Actual Rates, England^{3,10}

Figure 3. Past Projections of Life Expectancy at Birth and Actual Figures, England^{3,10}

Figure 4. Past Projections of Total Population Size and Actual Figures, England^{3,10}
Outline

- Motivation
- Overview of population projection methods
- **UN approach for probabilistic projections**
  - Probabilistic fertility projections
  - Probabilistic mortality projections
- UN probabilistic population projections
- Summary
- Software and reference
Probabilistic population projections: UN approach

- Probabilistic approach for modelling demographic transition using parametric functions (Raftery et al. 2012, 2014)

- Probabilistic projections of TFR from a time series model
  - Sample trajectories from the predictive distribution of future TFRs for each country and period
  - Convert each trajectory to age-specific fertility rates

- Currently, the UN uses a similar approach for life expectancy (probabilistic) but not migration (deterministic)

- Apply cohort-component projection model to each sample (Raftery et al. 2012; Gerland et al. 2014)

- Yields many possible futures of world population and thus probabilistic forecasts of any population indicator

- Method assessed by out-of-sample prediction for 5, 10, . . . , 30 years: Projection intervals reasonably well calibrated
Outline

- Motivation
- Overview of population projection methods
- UN approach for probabilistic projections
  - Probabilistic fertility projections
  - Probabilistic mortality projections
- UN probabilistic population projections
- Summary
- Software and reference
Probabilistic TFR projections

3 phases

Phase I: pre-transition high fertility

Phase II: fertility decline to below replacement level

Phase III: post-transition low fertility, with turnaround and fluctuations

Fertility transition has started in all countries ⇒ Phase I not modelled (all countries already in Phase II or III)
Phase II model: Fertility transition

- Fertility decline:
  - starts slowly at high TFR values
  - accelerates and peaks around TFR 5
  - decelerates towards the end of the transition
  - stops below replacement level
- TFR decline model: double logistic function (sum of 2 logistic curve)
- Random error term → Random walk with non-constant drift
Bayesian hierarchical model (BHM)

- Separate estimation for each country not feasible
  - Sparse data
  - Historical trend only partially observed

- Solution: For each country, borrow info from other countries

- Hierarchical model:
  - Country parameters distributed around world average
  - World and country parameters estimated simultaneously

- Between-country correlation in forecast errors included in prediction algorithm (Fosdick et al. 2014): Correlation is a function of whether 2 countries are neighbors, in the same UN region (out of 22), or had the same colonizer in 1945
India: Probabilistic TFR projection

- Double-logistic decline function
- Probabilistic projections of TFR
Uncertainty for high-fertility countries

- Uncertainty increases over time
- Uncertainty increases with current level of the TFR

**Burkina Faso:**
- 2005-10 = 6.1
- 2045-50 80% prediction interval = 2.2-4.3

**Mali:**
- 2005-10 = 6.8
- 2045-50 80% prediction interval = 2.6-5.8
Phase III: Post-transition low-fertility rebound

- Within observation period: Start of Phase III defined by the two earliest consecutive 5-year increases below 2
- Observed in 25 countries/areas: 20 European countries, plus Singapore, Hong Kong, USA, Canada, Barbados
Projections for low-fertility countries

- Bayesian hierarchical autoregressive (AR(1)) time series recovery model used for Phase III
- Use all below-replacement TFRs to estimate uncertainty in long-term projections
- Country-specific asymptotes estimated from data, resulting in global asymptote of 1.85 (80% projection interval 1.5-2.1)
Projections for countries with lowest fertility

- **Russian Federation**
- **China**
- **Singapore**
- **China, Hong Kong SAR**
Outline

- Motivation

- Overview of population projection methods

- UN approach for probabilistic projections
  - Probabilistic fertility projections
  - Probabilistic mortality projections

- UN probabilistic population projections

- Summary

- Software and reference
Projecting mortality

- Probabilistic projections of all future age-specific mortality rates desired for all countries.

- But data availability and quality vary greatly (WPP 2012):  
  - Good vital registration data: 91 countries (Germany)  
  - Incomplete vital registration data: 40 countries (Sri Lanka)  
  - Survey estimates of child and adult mortality: 61 countries (Senegal)  
  - Survey estimates of child mortality only: 17 countries (Laos)  
  - Limited or no data: 22 countries (North Korea)

- Estimate past life expectancy at birth ($e_0$) for all countries:  
  - Life tables (data for all ages, usually from VR)  
  - Model life tables (data for some ages, often from surveys)  
  - Life tables from similar countries (no data)

- Converts data from all countries to a common currency: $e_0$
Random walk with a Bayesian hierarchical model drift:
\[ l(c,t+1) = l(c,t) + g(\theta_c, l(c,t)) + \epsilon(c,t+1) \]

- Increase in \( e_0 \) modeled by BHM using double-logistic function
- Parameters estimated via MCMC
- Produce country-specific double-logistic parameters \( g(\theta_c,l(c,t)) \)
- BHM pools information about the rates of gains across countries
- Variance of distortion term, \( \epsilon(c,t) \), decreases as \( e_0 \) increases
Bayesian hierarchical model

- Separate estimation for each country not feasible because of few data and only part of the evolution observed
- Solution: For each country, draw on information from other countries
- Hierarchical model:
  - Double logistic parameters for a country distributed about “world average”
  - World and country-specific parameters estimated simultaneously
  - Based on observed gains in all countries, determine range of possible improvement curves
    - Get country-specific improvement curves
    - Combine overall outcome with observed decline in country
Probabilistic projection of mortality

- Project female $e_0$ using a similar BHM to TFR (Raftery, Chunn & Gerland 2013)
  - Asymptotic linear increase to 2100
  - Original choice: asymptote based on long-term trend in record $e_0$, or 2.3 years/decade (Oppen & Vaupel 2002)
  - Revised choice: asymptote based on trend in maximum age at death for Sweden since around 1970, or 1.3 years/decade (Wilmoth et al 2000, updated)

- Probabilistic projection of the female-male gap in $e_0$ (Raftery, Lalic & Gerland 2014)

- Convert each sample at each future year to age-specific mortality rates using a modified Lee-Carter method (essentially Lee-Miller)

- Kannisto function (logistic with upper asymptote of 1.0) used to extrapolate mortality rates to high ages (i.e., 100+)

- Result: Sample from predictive distribution of female and male age-specific mortality rates in each future time period and country
India: Probabilistic female $e_0$ projection

Double-logistic gain in $e_0$ function

Probabilistic projections of $e_0$
Japan: Probabilistic female e0 projection

Double-logistic gain in e0 function

Probabilistic projections of e0

5-year gains

Life expectancy at birth

-1 1 3 5 7

25 40 55 70 85 100 e(0)

1900 1960 2020 2080 Year

median
PI 80
PI 95

median
80% PI
95% PI
observed e0
Predictive distribution of gains in female $e_0$
Outline

- Motivation
- Overview of population projection methods
- UN approach for probabilistic projections
  - Probabilistic fertility projections
  - Probabilistic mortality projections
- UN probabilistic population projections
- Summary
- Software and reference
World population projections

Scenarios vs. probabilistic projections

80% and 95% prediction intervals
Where will the increase happen?

Africa:

- Now ~1 bil.
- 2100: 4.2 bil.
- 80% range: 3.5-5.1 bil.

Recent slowdown/stall of fertility decline in some countries of sub-Saharan Africa

- Unmet need for contraception (~25% in SSA over past 20 years)
- Ideal family size has declined but remains high (~4.5)
TFR projection for Nigeria

- Decline stalled past 10 years
- Expect renewed decline, but stall could continue
- Much uncertainty
- UN high/low scenarios too narrow
- 2100 median just above replacement: 2.2 (1.5-3.2)
Population projection for Nigeria

- Now: 160 mil.
- 2100: 914 mil.
  - Lower 10%: 532 mil.
- UN fertility variants (+/- half child) understate uncertainty of future trends (when today’s TFR > 3)
Brazil

Total fertility rate

Total population
What have we learned from probabilistic projections?

- UN fertility variants (+/- half child)
  - Overstate the uncertainty of future trends at the global level, and also for some low-fertility countries
  - Understate the uncertainty of future trends for high-fertility countries

- World population growth
  - 95% prediction interval for 2050: 9.0 - 10.1 billion
  - 95% prediction interval for 2100: 9.0 - 13.3 billion
  - Population stabilization unlikely in this century, but not impossible (probability ~30%)
What uncertainty is not (yet) accounted for?

- Uncertainty about the baseline population and current levels of fertility, mortality and migration
- Uncertainty about model specification (e.g., parameter to determine asymptotic increase of $e_0$)
- Uncertainty about future age patterns of fertility and mortality
- For countries with high prevalence of HIV, uncertainty about the future path of the epidemic
- Uncertainty about future sex ratios at birth
- Uncertainty about future trends in international migration
Uncertainty in past demographic estimates

Outline

• Motivation

• Overview of population projection methods

• UN approach for probabilistic projections
  • Probabilistic fertility projections
  • Probabilistic mortality projections

• UN probabilistic population projections

• Summary

• Software and reference
Key messages

- Population projections usually include a middle scenario taken as a “best guess” for future trends.
- Important to communicate that this “best guess” is only one possible outcome.
  - Any prediction of the future is uncertainty.
  - Smart policies should anticipate multiple possible outcomes.
- United Nations now employs two methods of illustrating the uncertainty of future trends:
  - Alternative scenarios.
  - Probabilistic models.
- Fertility variants (+/- half child) are useful illustrations but potentially misleading in some cases.
- Population stabilization is unlikely in this century.
Outline

• Motivation

• Overview of population projection methods

• UN approach for probabilistic projections
  • Probabilistic fertility projections
  • Probabilistic mortality projections

• UN probabilistic population projections

• Summary

• Software and reference
R packages (free open source) available at http://cran.r-project.org

- Probabilistic projections of total fertility rate: bayesTFR
- Probabilistic projections of life expectancy at birth: bayesLife
- Probabilistic population projections: bayesPop
- Graphical user interface: bayesDem, wppExplorer
R packages

- bayesDem
  - Estimate TFR parameters
  - Project TFR trajectories

- bayesTFR
  - Analyse and visualize MCMC results
  - Check convergence

- bayesLife
  - Estimate e0 parameters
  - Project female and male e0 trajectories

- bayesPop
  - Project population trajectories

UN datasets:
- wppExplorer
- wpp2012 (or 2010, 2008)
References

References

References


Acknowledgements

More than 8 years and ongoing of research and collaboration between the UN Population Division and Prof. Adrian Raftery (Department of Statistics of the University of Washington) and his team:

- All the team responsible (UN Population Division) for the 2012 revision of the World Population Prospects, especially Kirill Andreev, Thomas Buettner, Patrick Gerland, Danan Gu, Gerhard Heilig, Nan Li, Francois Pelletier and Thomas Spoorenberg

- Team members of the UW Probabilistic Population Projections (BayesPop) Project: Adrian Raftery, Leontine Alkema, Jennifer Chunn, Bailey Fosdick, Nevena Lalic, Jon Azose and Hana Ševčíková