

Reconstructing Past Populations With Uncertainty From Fragmentary Data

Mark C. WHELDON, Adrian E. RAFTERY, Samuel J. CLARK, and Patrick GERLAND

Current methods for reconstructing human populations of the past by age and sex are deterministic or do not formally account for measurement error. We propose a method for simultaneously estimating age-specific population counts, fertility rates, mortality rates, and net international migration flows from fragmentary data that incorporates measurement error. Inference is based on joint posterior probability distributions that yield fully probabilistic interval estimates. It is designed for the kind of data commonly collected in modern demographic surveys and censuses. Population dynamics over the period of reconstruction are modeled by embedding formal demographic accounting relationships in a Bayesian hierarchical model. Informative priors are specified for vital rates, migration rates, population counts at baseline, and their respective measurement error variances. We investigate calibration of central posterior marginal probability intervals by simulation and demonstrate the method by reconstructing the female population of Burkina Faso from 1960 to 2005. Supplementary materials for this article are available online and the method is implemented in the R package “popReconstruct.”

KEY WORDS: Bayesian hierarchical model; Cohort component model; Fertility rate; International migration; Markov chain Monte Carlo; Mortality rate.

1. INTRODUCTION

Every two years, the United Nations Population Division (UNPD) publishes detailed estimates and projections of key demographic quantities for all countries in the world, from 1950 to 2100; they appear in the *United Nations World Population Prospects* (UNWPP; United Nations 2009a, b). The parameters reported include age-specific fertility and mortality rates (vital rates), population counts, and international migration rates. These numbers are used for the development and assessment of policy and are of particular importance for countries that lack their own well-resourced official statistical systems. For many of these countries, the UNPD is a key partner in the process of compiling, analyzing, and publishing internationally comparable demographic data.

The UNWPP tables can be classified into two broad groups, known as estimates and projections. Projections are predictions about demographic parameters in the future while estimates concern populations of the past. We will use the term reconstruction, which we find less ambiguous. Nevertheless, use of the term “estimate” agrees with standard usage in statistics; data about demographic parameters in the past are used to estimate

the true values. Between the biennial editions of the UNWPP, new data are collected and old data become available, so revisions must be made to the reconstructions to accommodate it.

Demographers have long acknowledged the presence of uncertainty in their measurements and many so-called indirect methods exist for obtaining “best” point estimates (e.g., United Nations 1983). It is still common practice to produce ranges of projections of vital rates and population counts based on different scenarios; high-, medium-, and low-fertility, for example. These ranges cannot be interpreted probabilistically, but methods have been developed that do yield probabilistic projections; Lee (1998) and Booth (2006) provide reviews. In contrast, there has been relatively little work on the development of fully probabilistic methods for demographic reconstruction. Here we propose a general method for reconstruction that accounts for measurement uncertainty and works with the type of demographic data that have commonly been collected for most countries over the last 60 years or so.

The article is structured as follows. Existing methods of population reconstruction are reviewed in Section 2. In Section 3, we define our notation and parameters and describe the method. In Section 4, we investigate some statistical properties of our method through simulation before applying it to real data from Burkina Faso in Section 5. We close with a summary of the results and a discussion in Section 6.

2. EXISTING METHODS OF POPULATION RECONSTRUCTION

Outside of official statistical agencies, demographic reconstructions have been undertaken to study historical populations of the past (e.g., Wrigley and Schofield 1981) and to estimate excess mortality in crises such as famine or social upheaval (e.g., Boyle and Ó Gráda 1986; Heuveline 1998; Merli 1998; Goodkind and West 2001). The most commonly used methods are based on the demographic balancing equations. These are the

Mark C. Wheldon is Ph.D. Candidate, Department of Statistics, University of Washington, Seattle, WA 98195 (E-mail: mwheldon@uw.edu). Adrian E. Raftery is Professor of Statistics and Sociology, University of Washington, Seattle, WA 98195 (E-mail: raftery@uw.edu). Samuel J. Clark is Associate Professor of Sociology, University of Washington, Seattle, WA 98195, and Researcher at the Institute of Behavioral Science (IBS), University of Colorado at Boulder, CO, and at MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa (E-mail: samclark@uw.edu). Patrick Gerland is Population Affairs Officer, United Nations Department of Economic and Social Affairs, Population Division, New York, NY 10017 (E-mail: gerland@un.org). We thank the editor, associate editor, and three anonymous reviewers for very helpful comments. This work was supported by grants R01 HD054511, R01 HD070936, K01 HD057246, and 5R24HD042828 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD). The first author received partial support from a Shanahan Endowment Fellowship and NICHD training grant 5T32HD007543 to the Center for Studies in Demography & Ecology, and from the Center for Statistics and the Social Sciences, both at the University of Washington. The views and opinions expressed in this article are those of the authors and do not necessarily represent those of the United Nations. This article has not been formally edited and cleared by the United Nations.

basic accounting relationships, which state that the population size at time $t + \delta$ is equal to the size at time t plus births and immigrants, minus deaths and emigrants. These relationships are encoded in the cohort component method of population projection (CCMPP; Lewis 1942; Leslie 1945, 1948; Preston, Heuveline, and Guillot 2001). Given the size and age-structure of a population at some baseline time, its size and structure at any point in the future can be determined from the baseline population and the fertility, mortality, and international migration rates that prevail over the period of reconstruction.

The back projection method of reconstruction (Wrigley and Schofield 1981) attempts to apply these relationships in reverse by using an estimate of the population structure at the terminal year of the period of reconstruction. This approach is problematic since the CCMPP procedure is not formally invertible. To produce sensible results, some additional constraints have to be imposed or somewhat ad-hoc fixes applied. In response to these concerns, Lee (1971, 1974) proposed the method of inverse projection. Inverse projection enacts the reconstruction forward through time; it is named for the fact that, instead of estimating counts of births and deaths from rates, count data are used to infer rates. Both of these methods have been further developed since their inception. For example, McCaa and Barbi (2004) and Rosina (2004) described extensions of inverse projection, while Oeppen (1993) proposed generalized inverse projection (GIP). All of the extended methods remain purely deterministic.

A stochastic reconstruction method was proposed by Bertino and Sonnino (2003) who modeled childbirth and death as inhomogeneous Poisson processes. The method is designed to work with counts of births and deaths aggregated over age and so requires the analyst to specify model age-patterns for fertility and mortality, although different schedules can be chosen for different subperiods. These schedules are taken as the intensity functions of the process and realizations are simulated through time to produce sequences of empirical estimates of population age structures. Hence, the only source of variation accounted for is natural variation around the demographic rates; the total numbers of deaths by year are assumed to be recorded without error. Moreover, international migration is assumed to be negligible over the period of reconstruction.

Finally, all of the above methods were designed to work with individual data of the kind sometimes found in European parish registers, or aggregated summaries of them. The data available for many countries in our intended application are not as detailed. The aim of this article is to introduce a new method of reconstruction that uses measurements of demographic parameters from 50 to 60 years in the past from multiple noisy data sources, often available only for limited years or periods. Uncertainty due to measurement error is expressed as probability distributions and intervals rather than deterministic, scenario-based ranges. To our knowledge, there are no existing methods of human population reconstruction that have these features.

Daponte, Kadane, and Wolfson (1997) used an approach similar to ours to construct a counterfactual history of the Iraqi Kurdish population between 1977 and 1990. They, too, represented errors in the measurement of demographic parameters as the standard deviations of probability distributions. However, they used a low-dimensional parameterization of mortality and

fixed the age patterns of fertility. International and sub-national migration was assumed to be negligible.

Where migration is accounted for, it has been handled in different ways. Lee's inverse projection method estimates it as a residual if censuses are available at intermediate years in the reconstruction interval. Alho (1992) added it as extra error in measuring survival. Here, migration is treated explicitly and age-time specific estimates are available as for the other parameters.

Complex models have also been developed by fisheries researchers to study the population dynamics of marine life (e.g., Quinn and Deriso 1999). A large body of work also exists on the dynamics of land animal populations; Brooks, Catchpole, and Morgan (2000a) and Brooks et al. (2002) provide reviews. In most wildlife studies, data-collection methods and the type and amount of data available differ somewhat from those in human demography. Nevertheless, there are similar statistical challenges to overcome and these suggest a Bayesian approach may be appropriate. For instance, in ecology and demography, multiple sources of data informing the same demographic parameters often exist. Poole and Raftery (2000) found that Bayesian melding allowed these to be synthesized in a coherent manner (see also Givens, Raftery, and Zeh 1993; McAllister et al. 1994; Raftery, Givens, and Zeh 1995). Furthermore, population dynamics models are often over-parameterized (e.g., Lee 1985, 1993). The resulting ridges in the likelihood surface posed problems that Bayesian approaches have been able to overcome (e.g., Brooks et al. 2000b; Catchpole, Kgosi, and Morgan 2001).

Here we propose a Bayesian solution to the reconstruction problem in which the dominant source of uncertainty, measurement error, is adequately accounted for through fully probabilistic estimates. We handle multiple, noisy data sources and we account for uncertainty about migration as well as fertility and mortality. We do not require individual-level data of the European parish register kind.

3. METHOD

3.1 Notation and Parameters of Interest

In this article, we restrict our attention to the dynamics of populations of females only. The parameters of interest are age- and time-specific vital rates, net international migration flows, and population counts. We will refer to "international migration" as simply "migration" as this is the only type we consider. We use the symbols n , s , g , and f to denote population counts, survival (a measure of mortality), net migration (immigrants minus emigrants), and fertility, respectively. All of these parameters will be indexed by five-year increments of age, denoted by a , and time, denoted by t . Reconstruction will be done over the time interval $[t_0, T]$. The age scale runs from 0 to $A > 0$; in our application (Section 5) A is 80. To model fertility, we define $a_L^{[fert]} \leq a_U^{[fert]}$ where fertility is assumed to be zero at ages outside the range $[a_L^{[fert]}, a_U^{[fert]} + 5)$. Throughout, a prime indicates vector transpose.

Given $\mathbf{n}_{t_0} = (n_{0,t_0}, n_{5,t_0}, \dots, n_{A,t_0})'$, the vector of age-specific female population counts at baseline t_0 , and using " \circ " to denote entrywise product, the CCMPP gives $\mathbf{n}_{t+5} = \mathbf{Q}_t(\mathbf{n}_t + \mathbf{n}_t \circ \mathbf{g}_t/2) + \mathbf{n}_t \circ \mathbf{g}_t/2$, $t = t_0, t_0 + 5, \dots, T - 5$. Here, \mathbf{Q}_t is a K

by K matrix encoding fertility and mortality and K is the number of age groups. The quantity \mathbf{g}_t is the net number of migrants expressed as a proportion of the population size so that $\mathbf{n}_t \circ \mathbf{g}_t$ is the net number of migrants, a count. Projection proceeds in discrete steps; those alive at the beginning of each step are subjected to the fertility, mortality, and migration prevailing during the projection interval. This is a discrete time approximation to a continuous time process and there are standard adjustments to improve accuracy. Adding half of the migrants at the beginning of the projection step and the remainder at the end is one such adjustment. The specific form we use is:

$$\begin{bmatrix} n_{0,t+5} \\ n_{5,t+5} \\ \vdots \\ n_{A-5,t+5} \\ n_{A,t+5} \end{bmatrix} = \begin{bmatrix} \tilde{f}_{0,t} & \tilde{f}_{5,t} & \cdots & \tilde{f}_{A-5,t} & \tilde{f}_{A,t} \\ s_{5,t} & 0 & & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & s_{A,t} & s_{A+5,t} \end{bmatrix} \times \begin{bmatrix} n_{0,t} + g_{0,t}n_{0,t}/2 \\ n_{5,t} + g_{5,t}n_{5,t}/2 \\ \vdots \\ n_{A-5,t} + g_{A-5,t}n_{A-5,t}/2 \\ n_{A,t} + g_{A,t}n_{A,t}/2 \end{bmatrix} + \begin{bmatrix} g_{0,t}n_{0,t}/2 \\ g_{5,t}n_{5,t}/2 \\ \vdots \\ g_{A-5,t}n_{A-5,t}/2 \\ g_{A,t}n_{A,t} \end{bmatrix}, \quad (1)$$

where

$$\tilde{f}_{a,t} \equiv s_{0,t}(1 + \text{SRB})^{-1}(f_{a,t} + f_{a+5,t} \cdot s_{a+5,t})(5/2) \quad (2)$$

(Preston et al. 2001). Annualized, age-specific fertility rates enter Equation (1) through the transformation defined in Equation (2). This transformation maps the rates into proportions so that projection can be compactly expressed as a matrix multiplication. We use a tilde to represent this transformation and discuss it further below. SRB is the sex ratio at birth (the ratio of males to females among all births). Multiplication by SRB in Equation (2) ensures only the female births are kept. Throughout, we take SRB to be fixed at 1.05, a demographic convention. We will use $M(\cdot)$ to represent CCMPP and abbreviate Equation (1) as

$$\mathbf{n}_{t+5} = M(\mathbf{n}_t, \mathbf{f}_t, \mathbf{s}_t, \mathbf{g}_t). \quad (3)$$

The survival parameters, $s_{a,t}$, give the proportion of those aged $a - 5$ to a at time t who survive to be aged a to $a + 5$ at time $t + 5$, for $a = 0, 5, \dots, A + 5$. That is, $s_{a,t}$ is the proportion of people surviving into the age range $[a, a + 5)$ over the five years between t and $t + 5$. Also, $s_{A,t}$ is the proportion aged $[A - 5, A)$ at exact time t who survive to time $t + 5$, by which time they are in the age group $[A, \infty)$. We allow for subsequent survival in this age group by letting $s_{A+5,t}$ be the proportion aged $[A, \infty)$ at time t who survive five more years.

Age-specific mortality during a given time period will be summarized by life expectancy at birth (e_0), where

$$e_{0,t} \equiv 5 \sum_{a=0}^A \prod_{i=0}^a s_{i,t} + 5 \left(\prod_{i=0}^A s_{i,t} \right) (s_{A+5,t} / (1 - s_{A+5,t})). \quad (4)$$

The derivation is straightforward but requires additional concepts from demography (see the online supplementary materials). The quantity $e_{0,t}$ is the average age at death in a hypothetical cohort subjected for its entire life to the mortality conditions represented by the set of age-specific survival proportions, s_t .

The net number of migrants aged $[a, a + 5)$ to the population during the time period $[t, t + 5)$ is $g_{a,t}n_{a,t}$. The age-summarized measure of migration we use will be the average annual total net number of migrants that, for the period $[t, t + 5)$, is $(1/5) \sum_a g_{a,t}n_{a,t}$. We model $g_{a,t}$ instead of $g_{a,t}n_{a,t}$, however, because we expect variation over time and age in the former to be less dependent on magnitude.

The $\tilde{f}_{a,t}$ represent a combination of age-specific fertility and under-five mortality. When multiplied by $n_{a,t}$, they give the number of female births surviving to time $t + 5$. That is, the $\tilde{f}_{a,t}$ give the number of surviving female babies at $t + 5$ as a proportion of the number of females aged $[a, a + 5)$ alive at time t . We do not model $\tilde{f}_{a,t}$, however, because data are typically gathered to estimate fertility rates, $f_{a,t}$, not the proportions, $\tilde{f}_{a,t}$. Fertility rates are annualized occurrence/exposure rates. They give the ratio of births of both sexes to women aged a to $a + 5$ (occurrence) to person-years lived (exposure) during $[t, t + 5)$. Thus, the number of births in the projection step t to $t + 5$ is not quite $5f_{a,t}n_{a,t}$ because the denominator of $f_{a,t}$ is not $n_{a,t}$ (we need the factor of five because $f_{a,t}$ are single-year rates but the projection intervals are five years wide). The expression in Equation (2) is more accurate. To illustrate, calculating $n_{0,t+5}$ (and taking $g_{0,t}n_{0,t} = 0$ for clarity) gives

$$n_{0,t+5} = s_{0,t}(1 + \text{SRB})^{-1} \sum_{a=a_L^{[\text{fert}]}}^{a_U^{[\text{fert}]}} 5f_{a,t} \left(\frac{n_{a,t} + n_{a-5,t}s_{a,t}}{2} \right) \quad (5)$$

(where we use the SRB to keep only the female births). The quantity in the right-most parentheses in Equation (5) is a better approximation for the denominator of $f_{a,t}$ than $n_{a,t}$. We multiply by $s_{0,t}$ to account for mortality of births during the projection step.

Age-specific fertility rates for a given time period will be summarized by the total fertility rate (TFR) that, for the period $[t, t + 5)$, is defined as

$$\text{TFR}_t \equiv 5 \cdot \sum_{a=a_L^{[\text{fert}]}}^{a_U^{[\text{fert}]}} f_{a,t}. \quad (6)$$

It is the average number of children born to members of a hypothetical cohort of women who survive from age $a_L^{[\text{fert}]}$ to age $a_U^{[\text{fert}]} + 5$ and experience the age-specific fertility rates that were in effect during the period $[t, t + 5)$.

3.2 Data and Initial Estimates

Data on the parameters of interest come from numerous sources of varying quality and coverage. For example, estimates of fertility rates come from registers of births or from

surveys such as the Demographic and Health Surveys (DHSs), depending on the country. Moreover, several data points for the same age- and time-specific parameters are often available due to multiple overlapping surveys. The UNPD often adjusts these outputs to reduce systematic biases. They apply existing demographic techniques and draw on their expert knowledge about the specific data sources. Further details about the UNPD's methodology can be found in the literature by the United Nations (2010). In many cases, this bias-reduction stage is highly specific to country and parameter, and so we do not propose a general approach to replace this part of the analysis.

Instead, we take the outputs of this stage, which are single bias-reduced age-time series for all four parameters, as inputs to our model. We call the elements of these biased-reduced single series *initial estimates* and use an asterisk (*) to distinguish them from the true values. For example, we let $f_{a,t}$ be the true (unknown) fertility rate and $f_{a,t}^*$ be an initial estimate of it; similarly for the other parameters. We also use an asterisk to indicate estimates of population counts based on bias-reduced census counts (e.g., adjusted census counts based on post-enumeration surveys) or similarly adjusted counts from a comprehensive demographic survey. Hence, we let $n_{a,t}$ be the true population count and $n_{a,t}^*$ a census-based estimate of it.

3.3 Model Description

We will take t_0 and T to be, respectively, the earliest and most recent years for which census-based population counts are available. These will usually be bias-adjusted census counts, and we will just call them census counts. We will denote the years after t_0 for which census-based estimates of population counts are available by $t_0 < t_L^{[cen]}, \dots, t_U^{[cen]} \leq T$. Let $\theta = (n_{t_0}, f_{t_0}, f_{t_0+5}, \dots, f_T, s_{t_0}, \dots, s_T, g_{t_0}, \dots, g_T)$ be the vector of all age- and time-specific vital rates and migration parameters, as well as population counts at t_0 . These are precisely the inputs required by the CCMPP, $M(\cdot)$ in Equation (3). Census counts for years $t_L^{[cen]}, \dots, t_U^{[cen]}$ are not included in θ because they are not inputs to $M(\cdot)$.

Reconstruction is equivalent to estimation of θ . To achieve this, we propose the following hierarchical model that depends on the original data only through the initial estimates. We model the census counts after the earliest census at Level 1, conditional on the outputs of the CCMPP, which appear in Level 2. The initial estimates of fertility, mortality, and migration, as well as census counts at t_0 , are modeled at Level 3. Level 4 specifies informative prior distributions. We assume, a priori, that the elements of θ are mutually independent given θ^* .

$$\text{Level 1: } \log n_{a,t}^* | n_{a,t}, \sigma_n^2 \sim \text{Normal}(\log n_{a,t}, \sigma_n^2), \\ a = 0, 5, \dots, A \quad t = t_L^{[cen]}, \dots, t_U^{[cen]} \quad (7)$$

$$\text{Level 2: } n_{a,t} | n_{t-5}, f_{t-5}, s_{t-5}, g_{t-5} \\ = M(n_{t-5}, f_{t-5}, s_{t-5}, g_{t-5}), \\ a = 0, 5, \dots, A; t = t_0 + 5, t_0 + 10, \dots, T \quad (8)$$

$$\text{Level 3: } \log n_{a,t_0} | n_{a,t_0}^*, \sigma_n^2 \sim \text{Normal}(\log n_{a,t_0}^*, \sigma_n^2), \\ a = 0, 5, \dots, A \quad (9)$$

$$\log f_{a,t} | f_{a,t}^*, \sigma_f^2 \\ \sim \begin{cases} \text{Normal}(\log f_{a,t}^*, \sigma_f^2), & a = a_L^{[fert]}, \dots, a_U^{[fert]} \\ \text{undefined}, & \text{otherwise} \end{cases} \quad (10)$$

$$\logit s_{a,t} | s_{a,t}^*, \sigma_s^2 \sim \text{Normal}(\logit s_{a,t}^*, \sigma_s^2), \\ a = 0, 5, \dots, A + 5 \quad (11)$$

$$g_{a,t} | g_{a,t}^*, \sigma_g^2 \sim \text{Normal}(g_{a,t}^*, \sigma_g^2), \quad a = 0, 5, \dots, A, \quad (12)$$

(where $t = t_0, t_0 + 5, \dots, T$ in Equations (9)–(12))

$$\text{Level 4: } \sigma_v^2 \sim \text{InvGamma}(\alpha_v, \beta_v), \quad v \in \{n, f, s, g\}. \quad (13)$$

For $0 < x < 1$, $\logit x \equiv \log(x/(1-x))$. We also impose the restriction that all population counts be nonnegative, so that the joint prior on θ at time t is multiplied by

$$I(M(n_t, f_t, s_t, g_t) > 0) \equiv \begin{cases} 1 & \text{if, for all } a = 0, \dots, A, \\ & n_{a,t+5} \geq 0 \\ 0 & \text{otherwise.} \end{cases} \quad (14)$$

In standard Bayesian terms, Equation (7) is the likelihood of $n_{t_L^{[cen]}}, \dots, n_{t_U^{[cen]}}$, while Equation (13) is the prior distribution of σ_v^2 , specified by the user-defined hyperparameters α_v and β_v . Population counts at baseline are treated differently because Equation (8) is essentially a standard difference equation with initial condition θ , which contains n_{t_0} . These are the only population counts that $M(\cdot)$ takes as inputs. The remaining census counts, at times $t_L^{[cen]}, \dots, t_U^{[cen]}$, are compared with the outputs of $M(\cdot)$ via Equation (7). Inference will be based on the joint posterior distribution of θ .

The quantities involved, and their dependence relations, are summarized in Figure 1.

3.4 Determining the Hyperparameters

To determine plausible values of α_v and β_v , $v \in \{n, f, s, g\}$, we view the σ_v^2 as representing the variance of the errors in the initial estimates of the respective demographic parameters. Although prior knowledge about these variances is unlikely to be exact, we expect that informative estimates can be derived from experts' knowledge of the data sources. Methods for eliciting prior information from experts are numerous (e.g., O'Hagan et al. 2006). Here, we use a straightforward method based on the mean absolute error (MAE) of the transformed initial estimates.

Taking fertility rate as an example, note that Equation (10) implies that $\text{MAE}(\log f_{a,t} | \sigma_f^2) \equiv E(|\log f_{a,t} - \log f_{a,t}^*| | \sigma_f^2) = \sigma_f \sqrt{2/\pi}$. The prior distribution for σ_f^2 can be specified by choosing quantiles for $\text{MAE}(\log f_{a,t} | \sigma_f^2)$. Suppose expert opinion is that $\text{MAE}(\log f_{a,t} | \sigma_f^2)$ is likely to be close to 0.1, but could be as high as 0.5. This suggests setting $\text{median}(\sigma_f^2) = (0.1)^2 \pi / 2 = 0.016$ and the 0.975 quantile to $(0.5)^2 \pi / 2 = 0.393$. To find an inverse gamma distribution with these quantiles, we would fix α_f at a range of values between 0.3 and 6 and choose β_f such that $\text{median}(\text{MAE}(\log f_{a,t} | \sigma_f^2)) = 0.1$. The parameter α_f would then be chosen such that the 0.975 quantile of $\text{MAE}(\log f_{a,t} | \sigma_f^2)$ was about 0.5. This would give $\alpha_f = 1$ and $\beta_f = 0.0109$.

Since demographers are more used to thinking about untransformed fertility rates, it is useful to consider what specifying $\text{MAE}(\log f_{a,t} | \sigma_f^2)$ means on the original scale. MAE on the log scale approximates mean absolute relative error (MARE) on the original scale, where $\text{MARE}(f_{a,t} | \sigma_f^2) \equiv E(|f_{a,t} - f_{a,t}^*| | \sigma_f^2) / f_{a,t}^*$. This approximation is good for the MAE values used

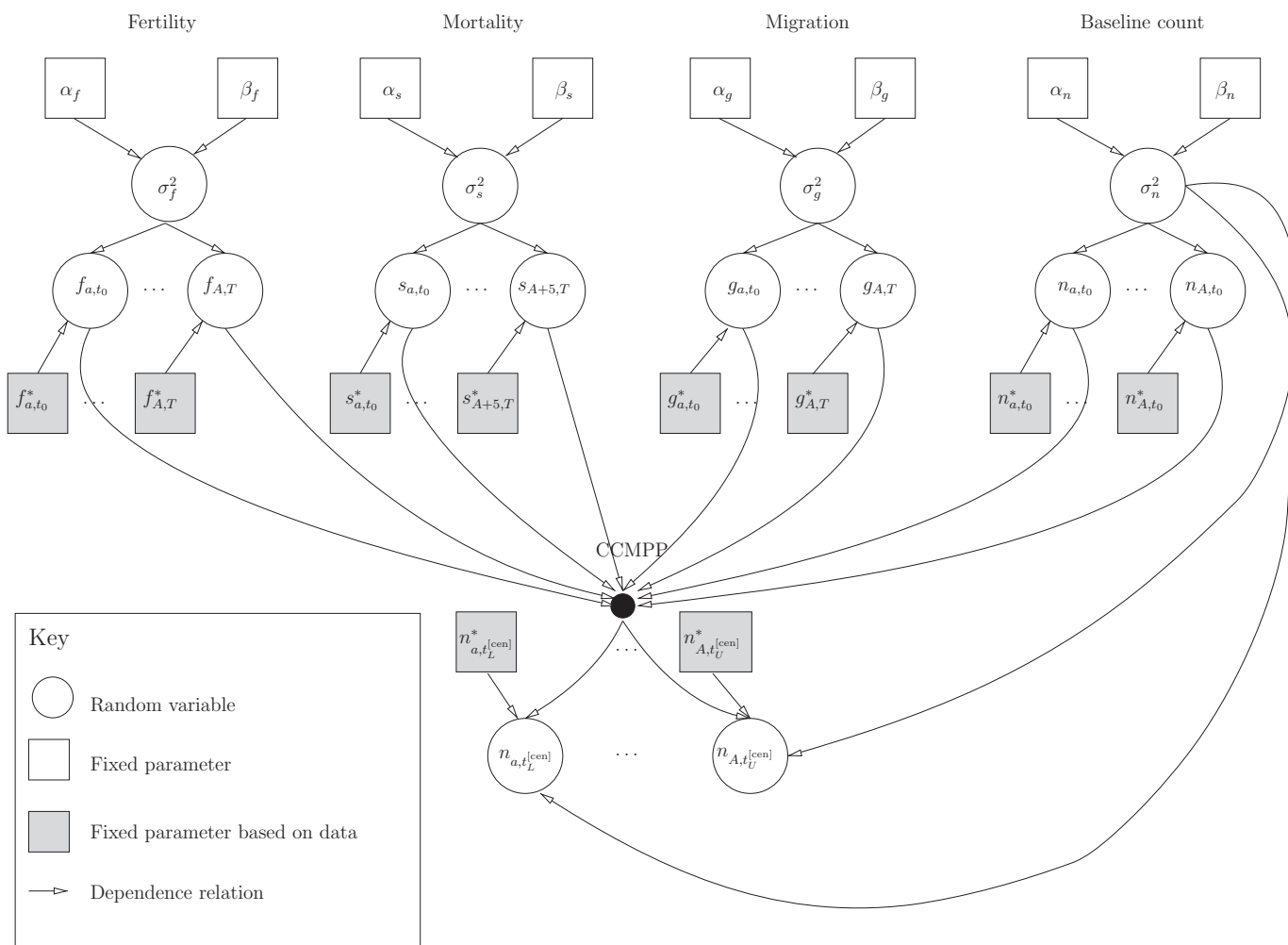


Figure 1. Plot showing the relationships between parameters, initial estimates, and the cohort component method of population projection in the estimation model.

here. The MARE has been used previously by demographers as a way of quantifying measurement accuracy (Keilman 1998). For the migration parameter, which is already a proportion, we specify the MAE directly.

The population count variance, σ_n^2 , is also modeled on the log scale and α_n , β_n are found in the same way as α_f and β_f . Survival is measured on the logit scale, but this should not be too difficult to interpret. Note that $E(|\logit \bar{s}_{a,t} - \logit \bar{s}_{a,t}^*| | \sigma_s^2) = E(|\logit s_{a,t} - \logit s_{a,t}^*| | \sigma_s^2)$ where $\bar{s}_{a,t} = 1 - s_{a,t}$. In practice, the $\bar{s}_{a,t}$ are close to zero, so that $\bar{s}_{a,t} \approx \bar{s}_{a,t}/(1 - \bar{s}_{a,t})$. Thus, specifying the MAE of the log-odds of survival is not that different from specifying it for the log-probability of death.

3.5 Estimation

We draw samples from the joint posterior using a Markov chain Monte Carlo (MCMC) sampler (Metropolis et al. 1953; Hastings 1970; Geman and Geman 1984). Without the restriction in Equation (14), the full conditional posterior distributions for the variance hyperparameters would be the usual conjugate inverse gamma distributions. With the restriction, the conjugate forms are not exactly correct but will probably be close to the true full conditionals. Therefore, to update these parameters we use the conjugate full conditional distributions as proposal

densities in Metropolis-Hastings steps. The posterior densities of the remaining parameters are not easy to express analytically since each vital rate enters the likelihood through the map $M(\cdot)$. Therefore, these parameters are updated using Metropolis-Hastings steps with univariate normal proposal densities, with variances tuned by the method proposed by Raftery and Lewis (1996). We will use the term “iteration” to refer to one complete sweep through all age- and time-specific parameters and variance parameters.

The R environment for statistical computing (R Development Core Team 2010) together with the CODA package (Plummer et al. 2006, 2010) were used for all data manipulation, model estimation, and output analysis. R code for the simulation study (Section 4), reconstruction of the female population of Burkina Faso, and model checks (Section 5) are available as online supplementary materials. The method is also implemented in the R package “popReconstruct.”

4. SIMULATION STUDY

We now describe the results of a simulation study carried out to investigate the calibration of posterior probability intervals for $f_{a,t}$, $s_{a,t}$, $g_{a,t}$, and n_{a,t_0} under repeated sampling of the initial estimates.

Table 1. True vital rates used in the simulation study

Age	Time period			
	(1960, 1965)	(1965, 1970)	(1970, 1975)	(1975, 1980)
Fertility rate				
0	0.00	0.00	0.00	0.00
5	0.40	0.40	0.40	0.40
10	0.30	0.30	0.30	0.30
15	0.00	0.00	0.00	0.00
Survival proportion				
0	0.90	0.90	0.90	0.90
5	0.95	0.95	0.95	0.95
10	0.85	0.85	0.85	0.85
15	0.80	0.80	0.80	0.80
20+	0.10	0.10	0.10	0.10
Migration proportion				
0	−0.03	−0.05	0.03	0.05
5	−0.05	−0.10	0.05	0.10
10	−0.06	−0.11	0.06	0.11
15	−0.01	−0.01	0.01	0.01

4.1 Inputs

The true vital and migration rates assumed to have prevailed in this population are shown in Tables 1 and 2. We denote these by $f_{a,t}^{[\text{true}]}$, $s_{a,t}^{[\text{true}]}$, $g_{a,t}^{[\text{true}]}$, and $n_{a,t_0}^{[\text{true}]}$. In practice, these would be the unknown, true values of the parameters $f_{a,t}$, $s_{a,t}$, $g_{a,t}$, n_{a,t_0} that appear in Equations (9)–(12). This is not intended to be a realistic model of a human population; datasets typically encountered in human demography have up to 18 age categories and any number of time periods. However, we believe that this reduced population model is of sufficient size and complexity to explore the characteristics of the statistical model while not being too computationally expensive. The TFR and e_0 were kept constant at 0.7 births per woman per year and 15.61 years, respectively, for the duration of the reconstruction. A varying pattern of migration was chosen that consisted of net out-migration in the first half of the reconstruction period followed by net in-migration in the second half. The magnitude of the flows was quite volatile, varying from 13% to 26% of the receiving population. Migration in both directions was concentrated in the two middle age groups.

The true population counts, denoted $n_{a,t}^{[\text{true}]}$, are shown in Table 2. Those for 1960 were chosen to represent a young population. Those for the subsequent time periods were derived by applying the CCMPP to the 1960 population using the vital rate and migration parameters in Table 1. Therefore, the underlying

Table 2. True population counts used in simulation study

Age	Year				
	1960	1965	1970	1975	1980
0	7500	8482	9453	11436	14504
5	6000	6886	7512	9280	11600
10	4000	4862	5293	6690	8651
15	3000	3404	3998	4762	6149

Table 3. Level 4 hyperparameters and selected implied quantiles of the Mean Absolute Error (MAE) of the respective demographic parameters used in the simulation study and the application to Burkina Faso. f and n are modeled on the log scale, s is modeled on the logit scale, and g is untransformed

v	α	β	Quantiles of $\text{MAE}(v)$				
			0.025	0.25	0.5	0.75	0.975
f, s, n	1	0.0109	0.0433	0.0707	0.1	0.1552	0.5232
g	1	0.0436	0.0867	0.1414	0.2	0.3104	1.0465

true population dynamics over the reconstruction period were completely and deterministically defined by Equation (1). The entries in Table 2 correspond to $n_{a,t}$, $t = t_0, t_L^{[\text{cen}]}, \dots, t_U^{[\text{cen}]}$ in Equation (8).

The hyperparameters of the inverse gamma distributions were determined as described in Section 3.4. The median MAEs for log-fertility rate, logit survival, and log-population counts were set to 0.1 with 0.975 quantiles of approximately 0.5. The same quantities for migration were set to 0.2 and approximately 1, respectively. The resulting values of α_v , β_v , and MAE quantiles are shown in Table 3. For fertility, $\text{MAE}(\log f_{a,t} \mid \sigma_f^2) = \sigma_f \sqrt{2/\pi}$; similarly for population count. For survival, $\text{MAE}(\text{logit } s_{a,t} \mid \sigma_s^2) = \sigma_s \sqrt{2/\pi}$ and for migration $\text{MAE}(g_{a,t} \mid \sigma_g^2) = \sigma_g \sqrt{2/\pi}$.

4.2 Study Design

The values in Tables 1–3 remained fixed throughout the simulation study and the initial estimates $f_{a,t}^*$, $s_{a,t}^*$, $g_{a,t}^*$, and $n_{a,t}^*$ were treated as random. We drew $n_{a,t}^* \mid \sigma_n^2$ from a $\log\text{Normal}(\log n_{a,t}^{[\text{true}]}, \sigma_n^2)$ distribution in accordance with Equation (7). The remaining initial estimates were drawn from distributions derived from Equations (9)–(12) in an analogous manner. For example, the $f_{a,t}^* \mid \sigma_f^2$ were drawn from a $\log\text{Normal}(\log f_{a,t}^{[\text{true}]}, \sigma_f^2)$ distribution.

The coverage of central marginal Bayesian confidence intervals (or credible intervals) under the model was estimated by the following experiment. For $j = 1, \dots, J$:

1. Randomly sample $\sigma_v^{2[j]}$, $v = n, f, s, g$, from Equation (13).
2. Generate initial estimates $f_{a,t}^{*[j]}$, $g_{a,t}^{*[j]}$, $n_{a,t}^{*[j]}$ for $a = 0, \dots, 15+$, $t = 1960, \dots, 1980$, $s_{a,t}^{*[j]}$ for $a = 0, \dots, 20+$ $t = 1960, \dots, 1980$, as described above.
3. Check that Equation (14) is satisfied by the initial estimates; if not return to Step 1.
4. Draw a large MCMC sample from the joint posterior and find the 0.025, 0.5, and 0.975 quantiles of the marginal distribution of each parameter.

The estimated coverage is then the proportion of the J Bayesian confidence intervals containing the known, true value for each parameter.

We set $J = 200$ and applied the estimation method described in Section 3.5. Initial values for the population counts, vital rates, and migration proportions were set to the initial estimates. Start

Table 4. Estimated coverage probabilities of 95% posterior intervals for all demographic parameters of interest

Years	Ages	Demographic parameter			
		Population	Fertility	Survival	Migration
1960	[0, 5)	0.930			
	[5, 10)	0.940			
	[10, 15)	0.945			
	[15, 20)	0.920			
	20+				
[1960, 1965)	[0, 5)			0.940	0.940
	[5, 10)			0.945	0.935
	[10, 15)		0.965	0.950	0.950
	[15, 20)		0.935	0.955	0.960
	20+			0.960	
[1965, 1970)	[0, 5)			0.950	0.940
	[5, 10)		0.920	0.945	0.940
	[10, 15)		0.960	0.935	0.965
	[15, 20)			0.925	0.955
	20+			0.935	
[1970, 1975)	[0, 5)			0.915	0.960
	[5, 10)		0.920	0.955	0.950
	[10, 15)		0.930	0.950	0.935
	[15, 20)			0.965	0.940
	20+			0.890	
[1975, 1980)	[0, 5)			0.950	0.930
	[5, 10)		0.920	0.950	0.970
	[10, 15)		0.955	0.945	0.950
	[15, 20)			0.950	0.955
	20+			0.975	

values for the variances were arbitrarily set to 5 as they appeared to have a negligible effect on the final results.

4.3 Results and Discussion

Point estimates of the coverage of the marginal 0.95 posterior probability intervals are shown in Table 4. These are all close to 0.95. In practical applications with real datasets, where the true parameter values are unknown, interest will be in interval estimates of the demographic parameters. These should be based on the joint posterior distribution. For illustration, we have plotted central marginal Bayesian confidence intervals of a selection of age-specific and age-summarized parameters that might be of interest based on the MCMC sample from a single replicate of the simulation study (Figure 2). For comparison, we have also plotted the true parameter values used throughout the simulation and the noisy initial estimates generated under the model.

Bayesian confidence intervals can be plotted for age-specific parameters as has been done for age-specific fertility rates in Figure 2(a). Confidence intervals for any function of the age-specific parameters can be obtained immediately by transforming each vector of age-specific values in the MCMC sample and computing the sample quantiles. Quantities of particular interest are the summary measures defined in Section 3.1. We show TFR, e_0 , and the average annual total net number of migrants in Figures 2(b)–(d).

5. RECONSTRUCTION OF THE POPULATION OF BURKINA FASO, 1960–2005

We now illustrate the method by reconstructing the female population of Burkina Faso from 1960 to 2005. Uncertainty in this case is nonnegligible due to the fragmentary nature of the available data. This application shows how our method is able to quantify this appropriately by producing probabilistic interval estimates.

5.1 Initial Estimates

Brief descriptions of our initial estimates are given below. Further details are in the online supplementary materials, including the initial estimates themselves. The parameters α_v and β_v , $v \in \{n, f, s, g\}$, were set to the same values as in the simulation study (Table 3); the MAEs given there were based on expert opinions provided by UNPD analysts for the case of Burkina Faso.

5.1.1 Population Counts. Population counts, $n_{a,t}^*$, in exact years 1960, 1975, 1985, 1995, and 2005 by sex were taken from the literature by the United Nations (2009a) that are based on a 1960–1961 demographic survey and censuses in 1975, 1985, 1996, and 2006. The United Nations (UN) figures are preferred over the raw census counts because important adjustments were made for underenumeration. This form of bias is more common in certain age-groups and efforts to reduce it are based on post-censal surveys.

Burkina Faso experienced a high level of population growth between 1960 and 2005; the total female population increased from 2.3 million to 6.9 million, an average of 2.4% per year. The population has a young age structure, as illustrated by the age-specific population counts in Figure 3.

5.1.2 Fertility Rates. Initial estimates of age-specific fertility rates, $f_{a,t}^*$, were derived from multiple, overlapping series of point estimates taken from the 1960 and 1991 demographic surveys, the 1976 census post-enumeration survey, the 1985, 1996, and 2006 censuses, and the 1992–93, 1998–99, and 2003 DHSs. Alkema et al. (2012) studied estimates of TFR using these data and found evidence of bias. Therefore, we took Alkema et al.'s (2012) median bias-adjusted estimates as our initial estimates of TFR and multiplied them by age-specific fertility patterns.

Age patterns sum to one and indicate the share of fertility attributable to each age-group. We obtained a separate pattern for each projection interval in the reconstruction period by smoothing the available point estimates over age, within interval, using *loess* (Cleveland 1979; Cleveland, Grosse, and Shyu 1992) and normalizing. The *loess* method performs a series of locally weighted regressions. Smoothing within five-year sub-interval (Figure 4) yielded trends that were also sensible, a priori, when viewed by five-year age group (see supplementary materials). No point estimates were available for the period 1965–1970. To generate initial estimates for this period, we multiplied the 1960–1965 age-pattern by Alkema et al.'s (2012) adjusted TFR estimate for 1965–1970.

5.1.3 Survival Proportions. Initial estimates of survival proportions, $s_{a,t}^*$, were derived from abridged demographic life tables constructed using data on under five and adult mortality from the 1960–1961 and 1991 national demographic surveys,

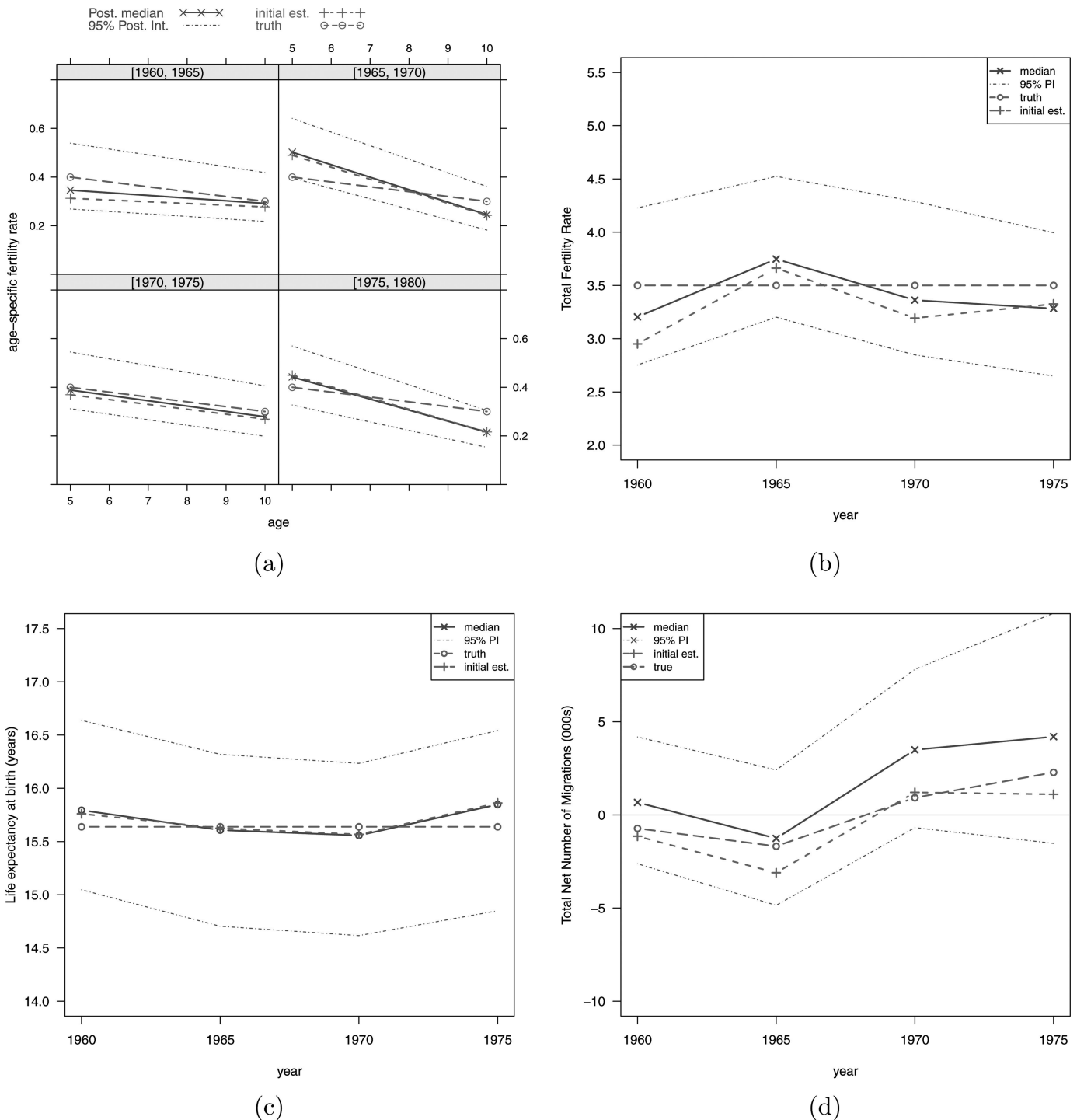


Figure 2. Ninety-five percent Bayesian confidence intervals for selected parameters from a single replication of the simulation study. (a) Age-specific fertility rate. (b) Total fertility rate. (c) Life expectancy at birth. (d) Average annual total net number of migrants. The online version of this figure is in color.

the 1992–1993, 1998–1999, and 2003 DHSs, and UNICEF’s Multiple Indicator Cluster Survey—Round 3 conducted in 2006. The tables were created by applying the Brass two-parameter relational logit model (Brass 1971) with the Timæus Sahelian standard (Timæus 1999). Standard methods were used to derive survival proportions from the life tables.

5.1.4 Migration Proportions. Estimates of migration of comparable detail to those for vital rates are seldom available for many countries. Some estimates for Burkina Faso by broad

age-group and sex are given by Condé (1980) for the period 1960–1975 who concluded that migration during this period was primarily labor migration. In addition, whole-population estimates for 1960–2005 are available from the literature by the United States Census Bureau (2008) and United Nations (2009a) that indicate sustained net out-migration over the period 1960–2000. The latter source suggests a reversal to net in-migration over 2000–2005.

We designed our initial estimates to reflect the direction and approximate magnitude suggested by these sources. The $g_{a,t}^*$

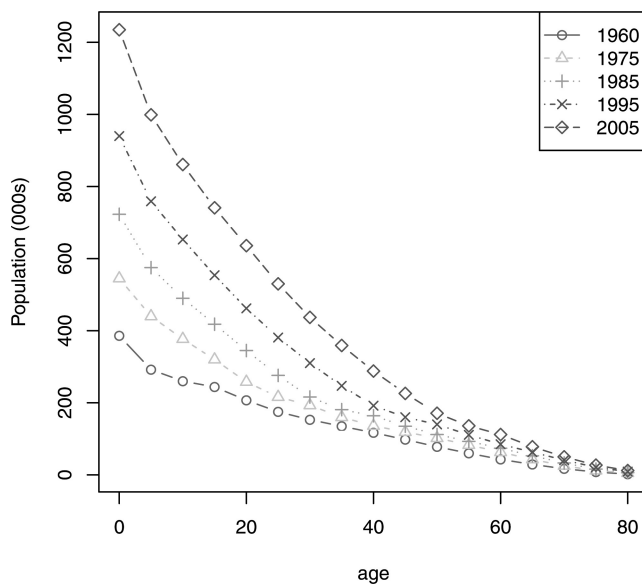


Figure 3. Population counts by five-year age group for the female population of Burkina Faso. The online version of this figure is in color.

were set to -0.055 for $15 \leq a \leq 50$, $1960 \leq t \leq 1995$, 0.055 for $15 \leq a \leq 50$, $t = 2000$, and zero otherwise. Thus, for example, our initial estimate for the net number of migrants over the five-year period 1960–1965 is centered at -5.5% of the 1960 population, an average of -1.1% of the 1960 population per year for the five years in the projection interval. The direction is reversed over 2000–2005, since Burkina Faso likely received refugees from the civil war that broke out in neighboring Côte

D'Ivoire during this period. The greater uncertainty about these initial estimates than about those for the other parameters is accounted for by setting α_g , β_g such that the median and the 0.975 quantile of the MAE were twice as large as the values used for the other parameters (Table 3).

5.2 Results

We summarize the full joint posterior distribution with posterior medians and 95% central Bayesian confidence intervals based on the marginal distributions of age-specific input parameters and age-summarized versions such as TFR and e_0 . We use half the width of the confidence intervals (“half-widths”) to indicate the magnitude of posterior uncertainty. Summarizing the joint posterior distribution of θ by quantiles of marginal distributions provides coherent and probabilistic interval estimates. They are coherent because uncertainty about all other parameters is accounted for. Our simulation study suggests that the intervals are also well calibrated in that they achieve their nominal coverage over repeated random sampling of the initial estimates and census data, under the model.

5.2.1 Population Counts. The posterior medians for population counts at baseline (Figure 5) are very close to their initial estimates; across age groups, the maximum absolute difference is 2.3% of the initial estimate. All of the intervals have half-widths less than 7.4% of the median, indicating that posterior uncertainty about this quantity is low.

5.2.2 Fertility. A similar plot for age-specific fertility rates is shown in Figure 6. Again, the posterior medians are close to their respective initial estimates, but the half-widths of the

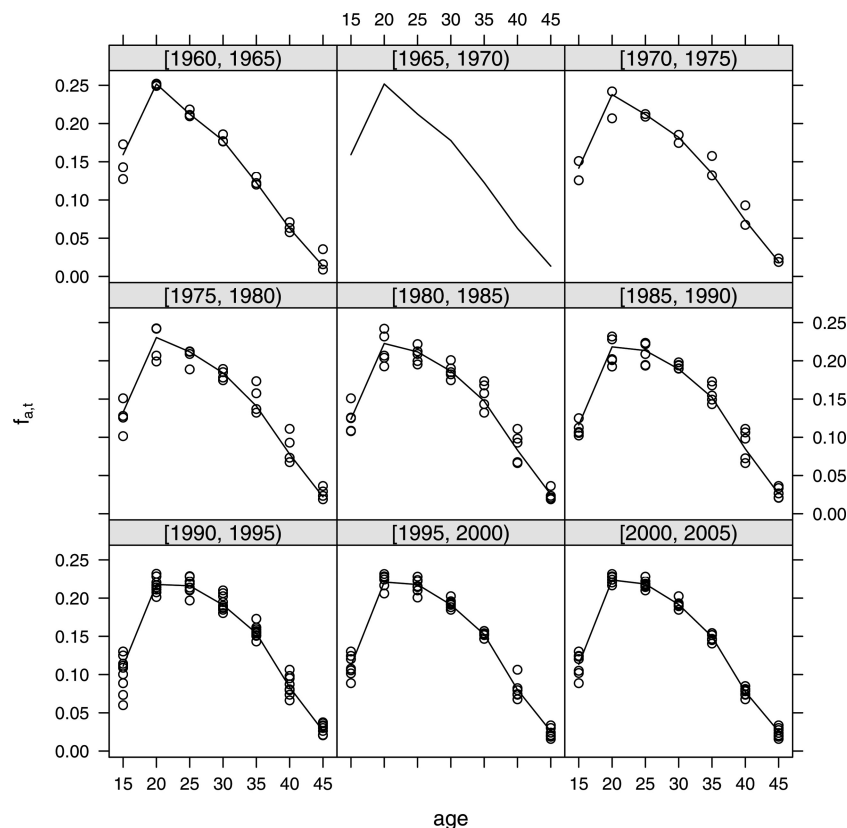


Figure 4. Data points for the initial estimates of age-specific fertility patterns of Burkina Faso women, 1960–2005, grouped by five-year sub-interval. The lines are the within-time loess smooths. The online version of this figure is in color.

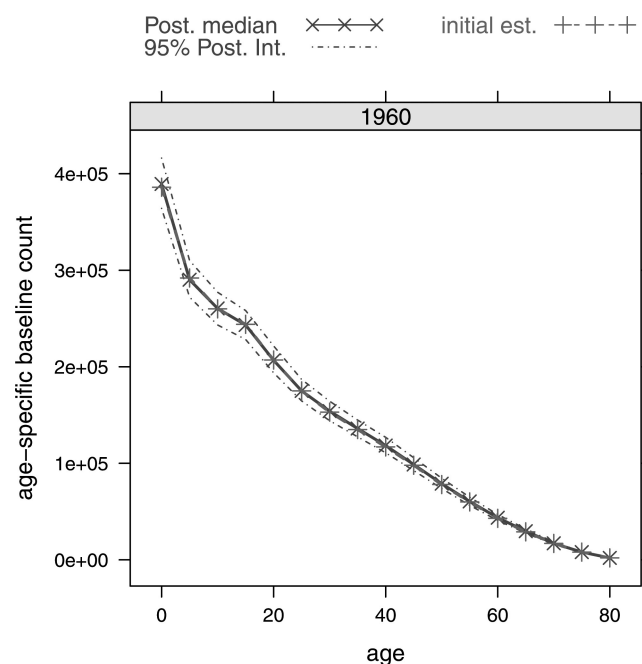


Figure 5. Ninety-five percent Bayesian confidence intervals for the female population of Burkina Faso by age in 1960. Also shown are the initial estimates. The online version of this figure is in color.

posterior intervals are wider, ranging from 18% to 21% of the median. This is because most of the information about fertility in the nonfertility parameters comes from the population

counts in the age range 0–5, and this depends mainly on the level of fertility, not how it is distributed across age group of mother. Information about the age pattern of fertility in our posterior distribution comes mostly from the data-derived initial estimates. The interval widths narrow as age-specific fertility approaches zero, which occurs at the extremes of the age range of nonzero fertility. Due mainly to biology, human fertility is known to be low at the extremes of this range with a high degree of certainty. The shape of our posterior intervals reflects this.

Posterior median estimates of TFR (Figure 7(a)) increased from 7.1 children per woman in 1960–1965 to about 7.4 over the period 1965–1980, and then decreased to 6.4 children per woman in 2000–2005. Over the entire reconstruction period, the limits of the 95% intervals are equivalent to about plus or minus half a child. The posterior medians are slightly lower than the initial estimates, which were based only on information about fertility collected mainly in surveys. Our posterior distribution for TFR also takes population counts, mortality, and migration into account. There is more information about TFR in these parameters than there is about the fertility age pattern, explaining why the interval half-widths for TFR are narrower, relatively, than those for the age-specific rates; all are less than 7.5% of the posterior median.

5.2.3 Mortality. Posterior estimates of mortality are presented in terms of e_0 and under-five mortality, $1 - s_{0,t}$ (Figures 7(c) and (d)). Under-five mortality is often viewed as an

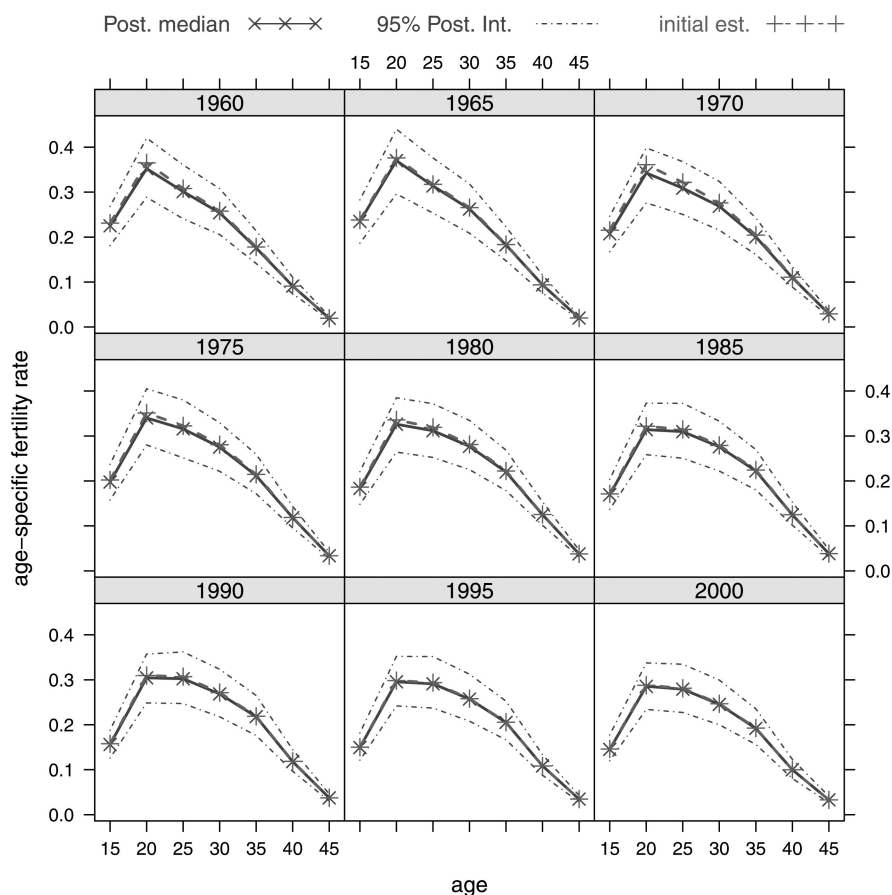


Figure 6. Ninety-five percent Bayesian confidence intervals for age-specific fertility rates for the female population of Burkina Faso, 1960–2005. Also shown are the initial estimates. The online version of this figure is in color.

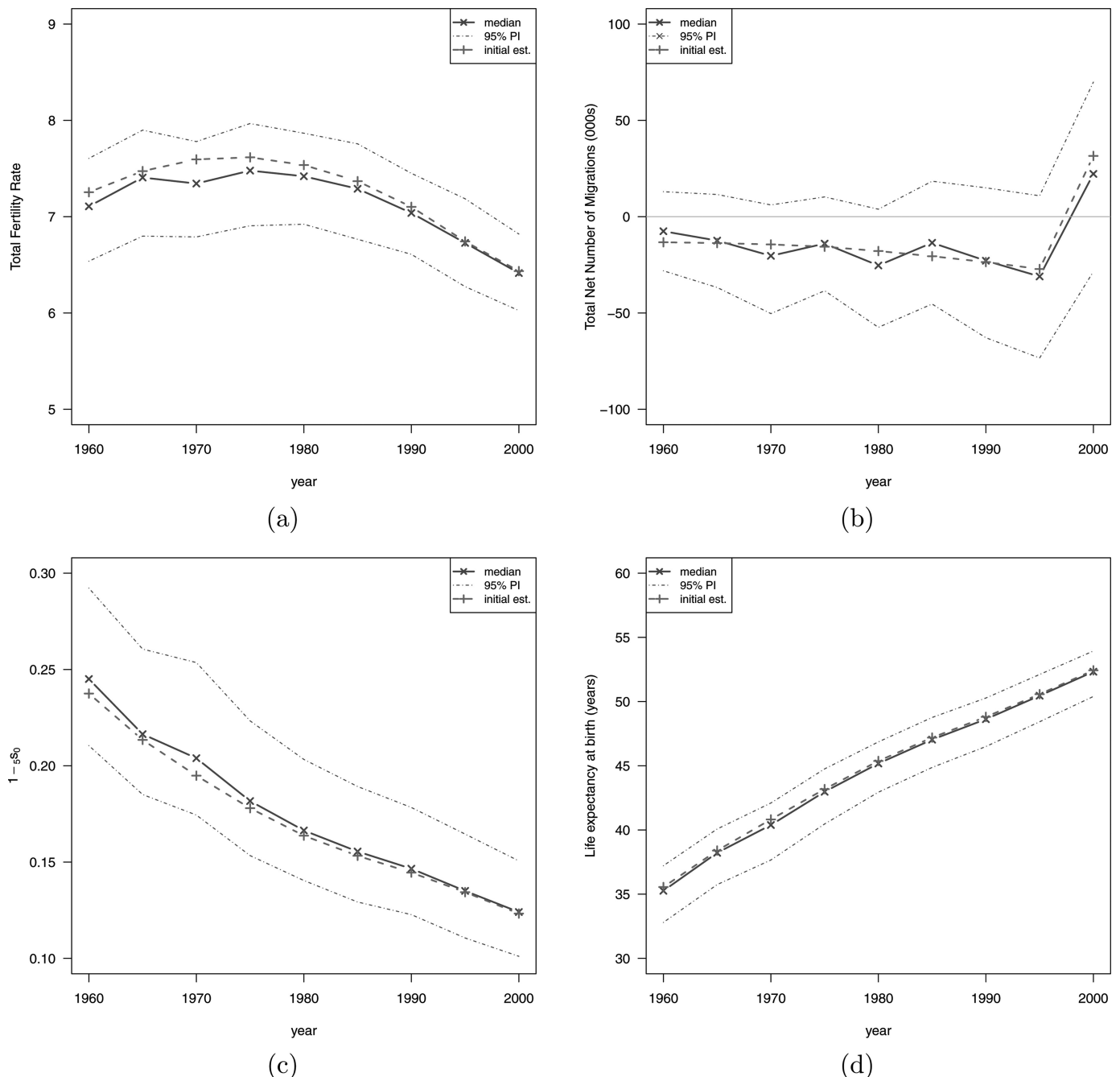


Figure 7. Ninety-five percent Bayesian confidence intervals for selected age-summarized parameters for the female population of Burkina Faso, 1960–2005. (a) Total fertility rate. (b) Average annual total net number of migrants. (c) Under-5 mortality. (d) Life expectancy at birth. The online version of this figure is in color.

indicator of a country's level of development. The posterior marginal distributions for these two parameters reflect a sustained improvement in mortality conditions in Burkina Faso over the period. Posterior median estimates of e_0 for 1960–1965 and 2000–2005 are 35 and 52 years, respectively. The interval half-widths decrease from 2.2 years in 1960–1965 to 1.8 years in 2000–2005, indicating a decrease in uncertainty about this parameter over the interval of reconstruction. Posterior median estimates of under-five mortality declined from 0.25 to 0.12 over the period (95% intervals: [0.21, 0.29] and [0.1, 0.15], respectively).

5.2.4 Migration. We summarize age-specific migration by the average annual net number of migrants added to the population, in units of 1000 (Figure 7(b)). The average of the posterior median estimates between 1960 and 2000 is -18 (thousand women) per year. Importantly, the 95% intervals contain zero at all time periods and, moreover, they are very wide. The interval for 1960–1965 is $[-28, 13]$, while it is $[-73, 11]$, for 1995–2000.

Burkina Faso has been characterized as a country of emigration between 1960 and 2000, with migration dominated by people moving to find work in neighboring countries. This view

is consistent with knowledge about the labor market during that period, as well as with data collected in neighboring countries (Condé 1980; United States Census Bureau 2008). These data are fragmentary, however, and there are no reliable estimates of the magnitude of the migratory flow. Our results reflect this situation. Our marginal posterior distribution is centered below zero, which suggests that, on balance, the available data are more consistent with a net outflow up to 2000. Nevertheless, there is insufficient information to rule out a zero, or even a small positive, net flow.

For 2000–2005, the 95% interval is $[-29, 70]$ with posterior median 22. While a positive net flow is suggested, the interval is very wide and contains zero. Clearly, there is a great deal of uncertainty about migration during this period, even after taking information about all the other parameters into account.

5.3 Model Checking and Sensitivity Analysis

The census counts, $n_{a,t}^*$, play a central role in our model; all vital rate and migration parameters are related to one another a posteriori through these counts. We conducted two checks; one to assess sensitivity to the form of the likelihood and another to assess out-of-sample predictive performance.

We assessed sensitivity to the use of the relatively light-tailed normal likelihood with constant variance using the approach by Carlin and Polson (1991). Replacing σ_n^2 with $\lambda_{a,t}\sigma_n^2$, $\lambda_{a,t} \sim \text{InvGamma}(1, 1)$, $t = 1960, 1965, \dots, 2000$, $a = 0, 5, \dots, 80+$, in Equations (7) and (9) relaxes the assumption of constant variance. Moreover, in comparison with Equation (7), the (marginal) likelihood under this formulation is the heavier-tailed Student's t (mean = $\log n_{a,t}$, scale = σ_n , df = 2) distribution. We refer to this modification as the “ t_2 likelihood model.”

Posterior marginal quantiles of the demographic parameters were examined to assess the need for this additional flexibility. We compared the empirical 0.025, 0.5, and 0.975 quantiles of the posterior marginal distributions of each age- and time-specific parameter using the MCMC samples from the runs under the modified and original models. This involved comparing two sets of multivariate distributions. Within each parameter, except migration, we calculated the absolute relative differences (ARDs) for each age and time, expressed as a percentage, and summarized them by their averages and maxima, taken over all ages and times. For example, the mean of the ARDs of the 0.025 quantiles for fertility rate was calculated as

$$\frac{100}{(17)(9)} \sum_a \sum_t \frac{|f_{a,t}^{[0.025, \text{orig}]} - f_{a,t}^{[0.025, t_2]}|}{f_{a,t}^{[0.025, \text{orig}]}} ,$$

where $f_{a,t}^{[0.025, \text{orig}]}$ and $f_{a,t}^{[0.025, t_2]}$ are the 0.025 quantiles of the posterior distribution for $f_{a,t}$ under the original and t_2 likelihood models. The numerals 17 and 9 correspond to the number of age groups and time periods, respectively. The maximum ARDs were computed similarly. Migration is expressed as a proportion so we calculated absolute (nonrelative) differences for this parameter.

Posterior distributions of the $\lambda_{a,t}$ were also examined to assess the need for the additional flexibility of the t_2 likelihood model. Posterior distributions of the $\lambda_{a,t}$ concentrated away from 1 would indicate that the original normal-based model fits poorly (Carlin and Louis 2009, chap. 4).

The predictive ability of the original model was tested by rerunning the analysis four times, each time omitting one of the census datasets. For $t = 1975, 1985, 1995, 2005$, we compared the posterior predictive distributions $n_{a,t} | n_{a,-t}^*$ with the point values $n_{a,t}^*$, where $n_{a,-t}^*$ is the set of census counts for age a for all years except t . We call this the “out-of-sample validation.” For each of the four runs, we summarized the difference between the posterior predictive median and the census counts using MARE, expressed as a percentage, where

$$\text{MARE} \equiv \frac{100}{(17)(4)} \sum_a \sum_{t=L}^{t_U^{[\text{cen}]}} \frac{|n_{a,t}^* - \tilde{n}_{a,t}|}{n_{a,t}^*},$$

$\tilde{n}_{a,t}$ is the sample median of the $n_{a,t}$ based on the MCMC sample and the numerals 17 and 4 refer to the number of age groups and censuses, respectively. Small values of MARE suggest that the model predicts the observed counts well.

5.3.1 Results. Quantiles of the posterior marginal distributions from the t_2 likelihood model were close to those from the original model (Table 5). All mean ARDs were below 6% and most maximum ARDs were below 5%. Large ARDs were found between the extreme quantiles of some population counts, the largest for the 0.975 quantile in 1960, age group 65–69 (ARD = 36%). However, on the raw scale, this represented a difference of less than 10,000 people, which is small relative to the total population size, the median estimate of which was slightly over two million in 1960 under both models.

The posterior distributions of the $\lambda_{a,t}$ were virtually all centered around 1. That of $\lambda_{80,1975}$ had the median furthest from 1, but was still very spread out; its 0.025, 0.25, 0.5, 0.75, and 0.975 quantiles were, respectively, 0.30, 0.93, 2.4, 9.4, and 167. This suggests that the extra flexibility of the t_2 likelihood model is not needed.

The MARE from the out-of-sample validation was 3.9%, indicating that the posterior predictive estimates of the census counts were close to the observed census counts.

Table 5. Mean and maximum Absolute Relative Differences (ARDs) between posterior marginal quantiles of the posterior distributions from the t_2 likelihood and original models. Absolute differences (ADs) are given for migration (see Section 5.3)

Parameter	Quantile	Mean ARD (%)	Max ARD (%)
Fertility rate	0.025	0.79	4.09
	0.500	0.22	0.65
	0.975	0.52	1.87
Survival proportion	0.025	0.38	2.67
	0.500	0.06	1.67
	0.975	0.30	2.39
Migration proportion	0.025	0.68	0.29
	0.500	0.72	2.36
	0.975	1.81	2.20
Population count, 1960	0.025	0.92	1.73
	0.500	0.20	0.48
	0.975	5.38	36.48
Population count, 1965–2005	0.025	0.92	19.08
	0.500	0.24	1.87
	0.975	1.07	24.36

Overall, these results indicate that the normal-based model as originally formulated in Section 3.3 fits the data reasonably well.

6. DISCUSSION

We have described a method for reconstructing past populations by age and sex that is designed to work with the type of data commonly collected in modern demographic surveys and censuses, especially in developing countries. Population dynamics are modeled by the well-known cohort-component method of population projection and measurement error is accounted for in a coherent, fully probabilistic manner through a Bayesian hierarchical model. Inference is based on the joint posterior distribution of all parameters, which depends on data through bias-reduced initial estimates. We applied our method to a real dataset and found that the widths of our posterior intervals indicated a nonnegligible amount of uncertainty about all parameters, but were of magnitudes within the range of variation demographers are used to when working with estimates and projections. The method is implemented in the R package “popReconstruct.”

Inferences are likely to be sensitive to changes in the initial estimates. However, these are primarily data-derived and our goal is to synthesize all of the available information; changes in these parameters should have an influence on the posteriors. In our application to the female population of Burkina Faso, the posterior medians were similar to the initial estimates in many cases. This suggests that the initial estimates of each parameter agreed closely with the census counts and the initial estimates of the other parameters. We would expect to see a greater difference between the posterior medians and initial estimates of parameters for which this agreement is weaker.

Lee (1971, 1974) and Oeppen (1993) proposed deterministic methods of population reconstruction. We have assumed a deterministic model only for the population dynamics. That is, given the true vital and migration rates, the evolution of the population is modeled deterministically. In contrast, Bertino and Sonnino (2003) gave a method in which the population dynamics were stochastic and the vital rates functioned as mean parameters. We agree with Pollard (1968; see also Cohen 2006) and expect that, in our applications, variation due to measurement error will overwhelm any additional variation arising from a stochastic population dynamics model. Moreover, Bertino and Sonnino’s (2003) method cannot be easily applied in our context because it was designed to take counts of baptisms and deaths by year as inputs. Information of this kind is seldom available in most of the cases where we wish to apply our method. Key features of our approach, therefore, are that the major source of uncertainty in population reconstruction, namely measurement error, is appropriately accounted for through fully probabilistic estimates and the method can be used with the kind of data available for most countries over the past 60 years.

We made the simplifying assumption of constant variance across age and time for each demographic parameter on the log or logit scale. This might seem restrictive; more complex variance structures for the other parameters could be proposed to allow for the fact that more is typically known about infant and child mortality than old-age mortality, for example. How-

ever, the benefits of this extension would have to be weighed against the additional complexity of having to estimate more parameters. As our checks indicated, using separate age- and time-specific variance parameters for population counts made little difference to the final results, probably because there was not enough information in the initial estimates to estimate them. The story might be different in a multicountry context, however. For instance, with demographically similar countries, it might be realistic to set some of the variance parameters constant across countries but allow for variation over time or age.

Census counts, or data of comparable reliability, are assumed available for the baseline year. This requirement could be removed by modifying Equation (9) so as to accommodate values of n_{a,t_0}^* derived from noncensus sources. Since these would probably be less reliable than census-derived estimates, this would mean replacing σ_n^2 in Equation (9) with a new variance parameter to account for the extra uncertainty. This could be modeled in the same way as the existing variance parameters per Equation (13). The period of reconstruction in our example was delimited by the years of the earliest and most recent data on population counts. This modification would permit the period of reconstruction to begin earlier if some initial estimate of population counts at an earlier baseline year were to be obtained.

No modifications are required to continue reconstruction beyond the year of the most recent census where initial estimates of vital rates and migration are available. However, the posterior distribution of the vital rates and migration for years beyond the most recent census will be based entirely on the initial estimates.

The structure of our model could be extended to produce more detailed reconstructions in several ways. An extension to two-sex populations is an obvious example. While we expect this to be feasible, dependencies between female- and male-specific parameters must be appropriately accounted for.

We used our method to reconstruct national populations. There are at least two ways in which subnational reconstructions could be obtained. The most straightforward would be to obtain sets of initial estimates for each subnational region of interest, and apply the model separately to each. Subnational estimates of population counts by age can be obtained from many censuses. In developed countries, vital registration systems would likely provide initial estimates of regional fertility and mortality. Less developed countries often lack these systems and such data might be difficult to obtain. Obtaining subnational estimates of net migration is likely to be difficult for most countries.

A more feasible approach is to use subnational initial estimates of population counts and set initial estimates of subnational vital rates to their national-level estimates. A minimal set of modifications to allow this could involve replacing the national-level population count vectors, \mathbf{n}_t , in Equation (1) with vectors of stacked sub-national counts. For R regions, the population count vector at time t , $\mathbf{n}_{:,t} \equiv (n_{0,t}, \dots, n_{A,t})'$, would be $(\mathbf{n}'_{:,1}, \dots, \mathbf{n}'_{:,R})'$. The projection matrix would take a block diagonal form with R blocks, one for each region. Similarly, the $g_{a,t}$ would be national-level initial estimates of migration proportions. Level 1, Equation (7), and Level 3, Equations (9)–(12), of the hierarchical model would be extended to include additional, similar terms for each region. For example, Equation (11) could be replaced by $\text{logit } s_{a,t,r} \mid s_{a,t}^*, \sigma_s^2 \sim \text{Normal}(\text{logit } s_{a,t}^*, \sigma_s^2)$ for regions

r, \dots, R . Level 4, Equation (13), would be unchanged for $v = f, s, g$ since we only have national-level initial estimates, but models with separate variance parameters for each region (i.e., $\sigma_{n,r}^2, r = 1, \dots, R$) would be worth investigating.

An advantage of this modification is that “hybrid” cases where subnational data are available for some time periods and/or parameters could be accommodated by using them to derive specific initial estimates and substituting these for the national-level estimates in the modified Level 3. Corresponding, additional variance parameters could also be added at Level 4. A possible shortcoming of this approach is lack of flexibility in cases where a few regional vital rates differ greatly from the national-level rates since the national level initial-estimates are used as fixed medians at Level 3. This might be detected by applying the approach by Carlin and Polson (1991) to the vital rate variance parameters, as was done in Section 5.3 for σ_n^2 . Alternatively, dependence among subnational vital rates could be explicitly modeled, perhaps using a spatial model in Level 3.

Migration is not split into its constituent inflows and outflows since the projection model (1) requires net migration as an input. Reliable data on migration over long periods are not usually available, especially for developing countries. Even where data on flows are collected, such as in the European Union, there can be significant disagreement in estimates of binational flows between the records of the sending and receiving countries (e.g., Raymer, de Beer, and van der Erf 2011). The UNPD has considered this issue and net migration is currently their preferred measure.

However, in cases where good information about both in- and out-migration is collected by means of population registers or border control agencies, it might be reasonable to model the two flows separately on a country-by-country basis. To allow this, the $g_{a,t}$ would be replaced with $g_{a,t}^I$ and $g_{a,t}^E$, say, for immigration and emigration, respectively. A minimal modification would then be to replace $g_{a,t}$ with $g_{a,t}^I - g_{a,t}^E$ in Equation (1) and replace Equation (12) with two similar terms for $g_{a,t}^I$ and $g_{a,t}^E$. If good information about the accuracy of the immigration and emigration data were available, then σ_g^2 could be replaced with two separate parameters; otherwise it would be unchanged.

SUPPLEMENTARY MATERIALS

The derivation of Equation (4); further information about the derivation of the initial estimates discussed in Section 5.1; the R code required to perform both the simulation study and the reconstruction of the female population of Burkina Faso (together with sensitivity analyzes); results of the Burkina Faso reconstruction as .csv_files.

[Received March 2012. Revised August 2012.]

REFERENCES

- Alho, J. M. (1992), “The Magnitude of Error Due to Different Vital Processes in Population Forecasts,” *International Journal of Forecasting*, 8, 301–314. [97]
- Alkema, L., Raftery, A. E., Gerland, P., Clark, S. J., and Pelletier, F. (2012), “Estimating Trends in the Total Fertility Rate With Uncertainty Using Imperfect Data: Examples From West Africa,” *Demographic Research*, 26, 331–362. [102]
- Bertino, S., and Sonnino, E. (2003), “The Stochastic Inverse Projection and the Population of Velletri (1590–1870),” *Mathematical Population Studies*, 10, 41–73. [97,108]
- Booth, H. (2006), “Demographic Forecasting: 1980 to 2005 In Review,” *International Journal of Forecasting*, 22, 547–581. [96]
- Boyle, P. P., and Ó Gráda, C. (1986), “Fertility Trends, Excess Mortality, and the Great Irish Famine,” *Demography*, 23, 543–562. [96]
- Brass, W. (1971), “On the Scale of Mortality,” in *Biological Aspects of Demography*, ed. W. Brass, London: Taylor & Francis. [103]
- Brooks, S. P., Catchpole, E. A., and Morgan, B. J. T. (2000a), “Bayesian Animal Survival Estimation,” *Statistical Science*, 15, 357–376. [97]
- Brooks, S. P., Catchpole, E. A., Morgan, B. J. T., and Barry, S. C. (2000b), “On the Bayesian Analysis of Ring-Recovery Data,” *Biometrics*, 56, 951–956. [97]
- Brooks, S. P., Catchpole, E. A., Morgan, B. J. T., and Harris, M. P. (2002), “Bayesian Methods for Analysing Ringing Data,” *Journal of Applied Statistics*, 29, 187–206. [97]
- Carlin, B. P., and Louis, T. A. (2009), *Bayesian Methods for Data Analysis* (3rd ed.), Boca Raton, FL: Chapman & Hall/CRC Press. [107]
- Carlin, B. P., and Polson, N. G. (1991), “Inference for Nonconjugate Bayesian Models Using the Gibbs Sampler,” *Canadian Journal of Statistics*, 19, 399–405. [107,109]
- Catchpole, E. A., Kgosi, P. M., and Morgan, B. J. T. (2001), “On the Near-Singularity of Models for Animal Recovery Data,” *Biometrics*, 57, 720–726. [97]
- Cleveland, W. S. (1979), “Robust Locally Weighted Regression and Smoothing Scatterplots,” *Journal of the American Statistical Association*, 74, 829–836. [102]
- Cleveland, W. S., Grosse, E., and Shyu, W. M. (1992), “Local Regression Models,” in *Statistical Models in S*, eds. J. Chambers and T. Hastie, Pacific Grove, CA: Wadsworth & Brooks/Cole, chap. 8. [102]
- Cohen, J. E. (2006), “Stochastic Demography,” *Encyclopedia of Statistical Sciences*, DOI: 10.1002/0471667196.ess2584.pub2. [108]
- Condé, J. (1980), “Migration in Upper Volta,” in *Demographic Aspects of Migration in West Africa* (vol. 2), eds. K. C. Zachariah, and J. Condé, Washington, DC: World Bank. [103,107]
- Daponte, B. O., Kadane, J. B., and Wolfson, L. J. (1997), “Bayesian Demography: Projecting the Iraqi Kurdish Population, 1977–1990,” *Journal of the American Statistical Association*, 92, 1256–1267. [97]
- Geman, S., and Geman, D. (1984), “Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images,” in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-6, pp. 721–741. [100]
- Givens, G. H., Raftery, A. E., and Zeh, J. E. (1993), “Benefits of a Bayesian Approach for Synthesizing Multiple Sources of Evidence and Uncertainty Linked by a Deterministic Model,” *Report of the International Whaling Commission*, 43, 495–500. [97]
- Goodkind, D., and West, L. (2001), “The North Korean Famine and Its Demographic Impact,” *Population and Development Review*, 27, 219–238. [96]
- Hastings, W. K. (1970), “Monte Carlo Sampling Methods Using Markov Chains and Their Applications,” *Biometrika*, 57, 97–109. [100]
- Heuveline, P. (1998), “Between One and Three Million: Towards the Demographic Reconstruction of a Decade of Cambodian History (1970–79),” *Population Studies*, 52, 49–65. [96]
- Keilman, N. (1998), “How Accurate are the United Nations World Population Projections?,” *Population and Development Review*, 24, 15–41. [100]
- Lee, R. D. (1971), *Econometric Studies of Topics in Demographic History*. Dissertation in Economics, Harvard University, Cambridge, MA. [97,108]
- (1974), “Estimating Series of Vital Rates and Age Structures From Baptisms and Burials: A New Technique, With Applications to Pre-Industrial England,” *Population Studies*, 28, 495–512. [97,108]
- (1985), “Inverse Projection and Back Projection: A Critical Appraisal, and Comparative Results for England, 1539 to 1871,” *Population Studies*, 39, 233–248. [97]
- (1993), “Inverse Projection and Demographic Fluctuations,” in *Old and New Methods in Historical Demography*, eds. D. S. Reher and R. Schofield, Oxford: Clarendon Press, pp. 7–28. [97]
- (1998), “Probabilistic Approaches to Population Forecasting,” *Population and Development Review*, 24, 156–190. [96]
- Leslie, P. H. (1945), “On the Use of Matrices in Certain Population Mathematics,” *Biometrika*, 33, 183–212. [97]
- (1948), “Some Further Notes on the Use of Matrices in Population Mathematics,” *Biometrika*, 35, 213–245. [97]
- Lewis, E. G. (1942), “On the Generation and Growth of a Population,” *Sankhyā: The Indian Journal of Statistics (1933–1960)*, 6, 93–96. [97]

- McAllister, M. K., Pikitch, E. K., Punt, A. E., and Hilborn, R. (1994), "A Bayesian Approach to Stock Assessment and Harvest Decisions Using the Sampling/Importance Resampling Algorithm," *Canadian Journal of Fisheries and Aquatic Sciences*, 51, 2673–2687. [97]
- McCaa, R., and Barbi, E. (2004), "Inverse Projection: Fine-Tuning and Expanding the Method," in *Inverse Projection Techniques: Old and New Approaches*, eds. E. Barbi, S. Bertino, and E. Sonnino, Berlin: Springer-Verlag, pp. 11–28. [97]
- Merli, M. G. (1998), "Mortality in Vietnam, 1979–1989," *Demography*, 35, 345–360. [96]
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953), "Equation of State Calculations by Fast Computing Machines," *Journal of Chemical Physics*, 21, 1087–1092. [100]
- Oeppen, J. (1993), "Back Projection and Inverse Projection: Members of a Wider Class of Constrained Projection Models," *Population Studies*, 47, 245–267. [97,108]
- O'Hagan, A., Buck, C. E., Deaneshkhah, A., Eiser, J. R., Garthwaite, P. H., Jenkinson, D. J., Oakley, J. E., and Rakow, T. (2006), *Uncertain Judgements: Eliciting Experts' Probabilities*, London: Wiley. [99]
- Plummer, M., Best, N., Cowles, K., and Vines, K. (2006), "CODA: Convergence Diagnosis and Output Analysis for MCMC," *R News*, 6, 7–11. [100]
- (2010), *CODA: Output Analysis and Diagnostics for MCMC* (Comprehensive R Archive Network, CRAN). Available at www.r-project.org. [100]
- Pollard, J. H. (1968), "A Note on Multi-Type Galton-Watson Processes With Random Branching Probabilities," *Biometrika*, 55, 589–590. [108]
- Poole, D., and Raftery, A. E. (2000), "Inference for Deterministic Simulation Models: The Bayesian Melding Approach," *Journal of the American Statistical Association*, 95, 1244–1255. [97]
- Preston, S. H., Heuveline, P., and Guillot, M. (2001), *Demography: Measuring and Modeling Population Processes*, Malden, MA: Blackwell. [97,98]
- Quinn, T. J., and Deriso, R. B. (1999), *Quantitative Fish Dynamics*, New York: Oxford University Press. [97]
- R Development Core Team (2010), *R: A Language and Environment for Statistical Computing*, Vienna: R Foundation for Statistical Computing. [100]
- Raftery, A. E., Givens, G. H., and Zeh, J. E. (1995), "Inference From a Deterministic Population Dynamics Model for Bowhead Whales," *Journal of the American Statistical Association*, 90, 402–416. [97]
- Raftery, A. E., and Lewis, S. M. (1996), "Implementing MCMC," in *Markov Chain Monte Carlo in Practice*, eds. W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, London: Chapman & Hall, pp. 115–130. [100]
- Raymer, J., de Beer, J., and van der Erf, R. (2011), "Putting the Pieces of the Puzzle Together: Age and Sex-Specific Estimates of Migration Amongst Countries in the EU/EFTA, 2002–2007," *European Journal of Population*, 27, 185–215. [109]
- Rosina, A. (2004), "Using Information on the Age Distribution of Deaths in Population Reconstruction: An Extension of Inverse Projection With Applications," in *Inverse Projection Techniques: Old and New Approaches*, eds. E. Barbi, S. Bertino, and E. Sonnino, Berlin: Springer-Verlag, pp. 29–38. [97]
- Timæus, I. M. (1999), "Notes on a Series of Life Table Estimates of Mortality in the Countries of the Sub-Saharan Africa Region," unpublished manuscript prepared for the World Health Organization. [103]
- United Nations (1983), *Manual X: Indirect Techniques for Demographic Estimation*, New York: Department of International, Economic and Social Affairs, United Nations. [96]
- (2009a), *World Population Prospects: The 2008 Revision* (Vol. I: Comparative Tables), New York: Department of Economic and Social Affairs, Population Division, United Nations. [96,102,103]
- (2009b), *World Population Prospects: The 2008 Revision* (Vol. II: Sex and Age Distribution), New York: Department of Economic and Social Affairs, Population Division, United Nations. [96]
- (2010), *World Population Prospects: The 2008 Revision* (Vol. III: Analytical Report), New York: Department of Economic and Social Affairs, Population Division, United Nations. [99]
- United States Census Bureau (2008), *International Data Base: Demographic Indicators, Burkina Faso, 1960–2000*. Available at <http://www.census.gov/population/international/data/idb/informationGateway.php>. [103,107]
- Wrigley, E. A., and Schofield, R. S. (1981), *The Population History of England, 1541–1871: A Reconstruction*, London: Edward Arnold. [96,97]