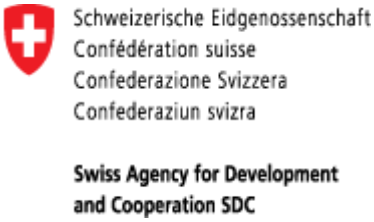
KNOMAD
Global Knowledge Partnership on Migration and Development

KNOMAD WORKING PAPER 30 - Annex

International Migration Projections: Methodology Brief

Thomas Buettner
Rainer Muenz

January 2018



The KNOMAD Working Paper Series disseminates work in progress under the Global Knowledge Partnership on Migration and Development (KNOMAD). A global hub of knowledge and policy expertise on migration and development, KNOMAD aims to create and synthesize multidisciplinary knowledge and evidence; generate a menu of policy options for migration policy makers; and provide technical assistance and capacity building for pilot projects, evaluation of policies, and data collection.

KNOMAD is supported by a multi-donor trust fund established by the World Bank. Germany's Federal Ministry of Economic Cooperation and Development (BMZ), Sweden's Ministry of Justice, Migration and Asylum Policy, and the Swiss Agency for Development and Cooperation (SDC) are the contributors to the trust fund.

The views expressed in this paper do not represent the views of the World Bank or the sponsoring organizations. Please cite the work as follows: Buettner, Thomas, and Rainer Muenz, 2017, International Migration Projections: Methodology Brief, KNOMAD Working Paper No.30, World Bank, Washington, DC.

All queries should be addressed to KNOMAD@worldbank.org. KNOMAD working papers and a host of other resources on migration are available at www.KNOMAD.org.

International Migration Projections: Methodology Brief*

Thomas Buettner and Rainer Muenz†

Abstract

This paper documents key elements and their mathematical representation of a projection methodology for international migration. The methodology rests on the well-established multiregional model in demography, but was reformulated to allow for more realistic migration assumptions that depend on sending and receiving countries. The paper develops the methodology in a consistent and transparent way in matrix notation. The starting point is the generating matrix of multiregional models, followed by the multiregional life table and corresponding projection model. The paper distinguishes between two types of multiregional projection models—the standard model, which implements migration solely in terms of emigration intensities, and a model that allows for interactions between the regions or countries involved. The second projection model is nonlinear and requires an iterative approach. The paper implements procedures to estimate the demographic events that are associated with the multiregional model. In the classical formulation of the multiregional projection model, births are (implicitly) covered, but deaths and migrants by age are neglected. The explicit inclusion of all vital events adds considerably more analytical power to multiregional population projections. The last addition to the multiregional methodology is the replacement of the common assumption of a uniform distribution of (decrement) events within age groups by a more realistic assumption that rests on approximations by local fits and an empirical model especially for the first age group. The methodology was implemented in a multiregional analysis and projection software package and coded in the high-level statistical language R.

See Working Paper 30: Modeling Alternative Projections of International Migration here.

Key words: Multiregional methodology, International migration, Population projections, Matrix models, Multiregional vital events, Non-linear decrements

---

* This paper was produced for KNOMAD's Thematic Working Group (TWG) on Data on Migration and Demographic Changes. KNOMAD is headed by Dilip Ratha; the Data on Migration and Demographic Changes TWG is chaired by Rainer Muenz; and the focal point in the KNOMAD Secretariat is Supriyo De. Comments by two anonymous reviewers are gratefully acknowledged, as is the meticulous editing work by Sandra Gain.

†Thomas Buettner is a demographer who was working at the United Nations Population Division before his retirement. Rainer Muenz is Advisor on Migration and Demography at the European Political Strategy Centre, European Commission. The authors may be contacted at planetbuettner@gmail.com; rainer.munz@ec.europa.eu.
The analysis and proposals expressed in this paper are the authors' personal views and do not represent the positions of their current or former employers.

Table of Contents

1. Multiregional Demographic Models 1

1.1. Objective and Organization of the Paper 1
1.2. The Generating Matrix of Multiregional Models 1
1.3. Multiregional Life Table 3
1.4. Multiregional Population Projections I 8
1.5. Multiregional Population Projections II 10

2 Events in a Multiregional Model 12

2.1. Demographic Events Gap 12
2.2. Births 12
2.3. Deaths 13
2.4. Migrants 14

3. Accounting for Nonlinear Decrements 15

3.1. From Rates to Probabilities 15
3.2. First Age Group 16
3.3. Middle Age Groups 16
3.4. Last, Open-Ended Age Group 17

Annex A. Notation and Symbols 18

A.1. Notation 18
A.2. List of Symbols 19

Bibliography 22

Tables

Table 1: Basic Demographic Rates 4
Table 2: Function Parameters for Estimating the Mean Duration of Transfer for Age Group (0,5) 16

Box

Box 1: Column-Row versus Row-Column Notation 3

1. Multiregional Demographic Models

1.1. Objective and Organization of the Paper

This paper documents the methodology and a preliminary software implementation of a novel approach for a more realistic projection of international migration by extending and refining the well tested demographic multiregional demographic model. The two most important additions to the classic multiregional methodology presented here are:

1. The introduction of (international) migration as the interaction between sending and receiving countries (see Buettner and Muenz 2017) into the multiregional projection model. This requires moving beyond net migration as the dominant concept in international population projections and implementing (international) migration as flows/transitions. The groundbreaking work of Abel (2009, 2013) and Abel and Sanders (2014a) opened the door for this endeavor on the global scale. It is now possible to replace the emigration-based conceptualization of migration in multiregional demographic models by introducing the receiving country as an additional factor determining the level of international migration flows.
2. The addition of methodology to derive explicitly the hitherto neglected demographic events of deaths and migration/migrants. This addition is important for the detailed and transparent analysis of projected migration streams.

The paper is organized as follows. This section sets out the mathematical basis for the multiregional methodology and population projections, including a more realistic handling of bilateral migration. Section 2 introduces additions and extensions to that methodology for the full inclusion of demographic events in multiregional population projections. Section 3 describes a practical way to consider the nonlinearity of decrement events in the model. The annex explains the notation and provides a list of symbols used in the paper.

1.2. The Generating Matrix of Multiregional Models

Traditionally, demographic analysis was mostly concerned with births and deaths; migration was absent from many theoretical models. The consequences were that demographic dynamics occurred—figuratively speaking—on an isolated island. People were born, aged, and reproduced and, finally, died. Leaving that island, or entering it, was not considered.

The multiregional model[1] is concerned with the movement of people during their life course between different geographic locations (or other states, such as marital status, health, and so forth) while exposed to basic demographic life events, such as births and deaths. The most important insight behind the multiregional model is the broadening of the concept of leaving the system of regions considered. Traditionally, the basic event of leaving the system was death—moving from being alive to the absorbing state of death, from which return is not possible. The multiregional model defines the movement out of

[1] Multiregional models have been generalized to cover other demographic dimensions and their associated events of change, such labor force participation, marital status, health status, and so forth. To reflect this much broader area of analysis, the term multistate models was introduced. As this paper is concerned with countries or aggregates of countries, the term multiregional model is used throughout.

a geographic entity or region as another event of leaving, here, the current region of residence. Unlike the absorbing state of death, further movements, including return movements, are (in principle) allowed.

The multiregional demographic model is defined for a system of *m* regions. In its most basic form, the model is arranged in an *m*-by-*m* matrix, $\mathbf{M}$. The regions may be a set of individual countries or geographic or economic aggregates of regions or subnational regions.

$$\mathbf{M}(x,n) = \begin{bmatrix} \left[d_1(x,n) + \sum_{j \neq 1} e_{1j}(x,n) \right] & -e_{21}(x,n) & \cdots & -e_{m1}(x,n) \\ -e_{12}(x,n) & \left[d_2(x,n) + \sum_{j \neq 2} e_{2j}(x,n) \right] & \cdots & -e_{m2}(x,n) \\ \vdots & \vdots & \vdots & \vdots \\ -e_{1m}(x,n) & -e_{2m}(x,n) & \cdots & \left[d_m(x) + \sum_{j \neq m} e_{mj}(x,n) \right] \end{bmatrix} \quad (1)$$

Matrix **M** is in a column-row format, with the columns representing the region of origin and the rows the destination (box 1).

In this arrangement, all decrement events (death and emigration) attributed to region i are related to this region's population,[2] and the corresponding death rates and emigration rates are classical exposure or occurrence rates. The main diagonal in matrix $\mathbf{M}$ contains what is called the total rates of exit of a region, for example, the death rate of that region and the sum of all emigration rates $d_i(x,n) + \sum_{j \neq i} e_{ij}(x,n)$. The emigration rates from region i to j are placed in the off-diagonal locations in column *i*.[3]

It is easy to see that the column sum of region i equals the death rate of that region. As the total rate of exit on the main diagonal combines death and emigration, adding the negative values of the emigration rates yields the original death rate. A similar property occurs in the matrix of survivorship proportions, which allows the "recovery" of deaths and movements/emigration in the projection process (equation 16).

Each region is (potentially) also a receiver of migrants. The inflow of immigrants into region i is specified in terms of emigration rates from regions $j \neq i$, which are placed on the rows of matrix $\mathbf{M}$. It is important that the events of migratory moves are exclusively specified as emigration rates.

[2] The population exposed to the decrement events is approximated by an estimate of the person-years lived during the period, here written as $\overline{K}_i$.

[3] The total rates of exits/decrements and their migratory components are later shown to be instrumental for the recovery of the events of deaths and migratory moves in the multiregional model.

Box 1: Column-Row versus Row-Column Notation

The tabular arrangement of bilateral migration data implies direction between origin and destination. In multiregional demography, there are two different but equivalent notations.

For the formulation of the multiregional modeling framework, Rogers uses a *column-row format* (denoted the Rogers format henceforth; see Rogers and Willekens (1976)). In the column-row format, columns are the origins and rows are the destinations. Hence, emigrants move along columns. The United Nations data on migrant stocks (UNPD 2015b) are in column-row format. Caswell (2001) also employs the column-row format.

The conventional form of matrix notation, used, inter alia, by Schoen (1988, 2006), is the *row-column* format. In the row-column format, the rows are the origins and the columns are the destinations. Emigrants move along rows. Abel (2013) presents his stock-to-flow methodology as well as the resulting data sets in row-column format.

The specific arrangement of the migration elements is of critical importance in multiregional demography, as it determines the order of matrix operations, since matrix multiplication is not commutative.

Matrix $\mathbf{M}$ is the generating matrix[4] of multiregional models; it is the basis of the multiregional life table, which is introduced next. The multiregional projection model (sections 1.4 and 1.5), in turn, is based on the multiregional life table.

1.3. Multiregional Life Table

The multiregional model is based on observed or estimated demographic occurrence/exposure rates that capture the dynamics of the demographic system under review. To calculate the occurrence/exposure rates of these events, the respective events are related to the population exposed to the risk of experiencing the event, expressed as the number of person-years lived during the interval/period and in a certain age group (table 1).

Assuming a linear distribution of decrements, the population at risk may be approximated by the arithmetic average of the population at the beginning and end of the period.[5]

$$\bar{K}_i^{t,n}(x,n) \approx \frac{n}{2}\left[K_i^t(x,n) + K_i^{t+n}(x,n) \right] \tag{2}$$

[4] The term "generating matrix" is inspired by Willekens (2008, 129), who uses it in a probabilistic context. Schoen (2006) uses the term "multistate data rate matrix."

[5] Assuming a linear approximation, in the case of person-years lived, is an acceptable approach. Other approximations are only marginally more accurate. The assumption of exponential growth of the population between two points in time, for instance, would be an alternative to the linear approximation (Hill 1990, 60), but special arrangements would need to be made for the case of a zero-growth rate.

Table 1: Basic Demographic Rates

Rate	Births	Deaths	Emigration
Crude rate	$b_i = \frac{B_i}{\bar{K}_i}$	$d_i = \frac{D_i}{\bar{K}_i}$	$e_{ij} = \frac{E_{ij}}{\bar{K}_i}$ (3)
Age-specific rate	$b_i(x,n) = \frac{B_i(x,n)}{\bar{K}_i(x,n)}$	$d_i(x,n) = \frac{D_i(x,n)}{\bar{K}_i(x,n)}$	$e_{ij}(x,n) = \frac{E_{ij}(x,n)}{\bar{K}_i(x,n)}$ (4)
Range of index *i*	$i = 1,\cdots,m$	$i = 1,\cdots,m$	$i = 1,\cdots,m;\ j \neq i$

The events of change as summarized in generating matrix $\mathbf{M}$ are used to calculate the corresponding indicators of a multiregional life table. The multiregional life table then provides the indicators needed for the cohort-component methods of multiregional population projections. In contrast to the classical cohort-component projection method for single populations open to (net) migration, the multiregional model includes migratory movements implicitly and consistently.

The first step is to transform the observed age-specific rates in matrix $\mathbf{M}$ into probabilities of surviving from exact age x to exact age $x+n$, assembled in matrix $\mathbf{P}(x)$. The transformation of rates into corresponding probabilities is a critical step in constructing a (multiregional) life table. The most common approach to transform exposure/occurrence rates into age-specific probabilities is to assume a uniform distribution of events, which leads to an easy numerical implementation (equation 5):

$$\mathbf{P}(x) \approx \left[\mathbf{I} + \frac{n}{2}\mathbf{M}(x,n)\right]^{-1} \left[\mathbf{I} - \frac{n}{2}\mathbf{M}(x,n)\right] \tag{5}$$

The assumption of a uniform distribution of decrements is simple and intuitive, but crude. It has the limitation that mortality rates or exit rates cannot exceed the value of 0.4 in the case of five-year age groups; otherwise, implausible negative probabilities are obtained.[6]

There are several solutions to account for the nonlinearity of the decrements, especially for age groups of more than one year in width. A simple and tested assumption is to introduce a correction factor, commonly written ${}_n a_x$, that reflects the of number of years lived in the interval by those who die in it (Keyfitz 1968, 12). A more general term that reflects the multiregional case where each region may be subject to multiple decrements is the Mean Duration at Transfer (MDT) measure; see Schoen (1979; 1988, 14–15; 2006, 18). To distinguish between the single decrement and the multiregional case, in this paper, the symbol ${}_n a_x$ is replaced by ${}_n \delta_x$ and Δ is the equivalent symbol for the multiregional case.

The values of ${}_n \delta_x$ or the corresponding matrix expression $\mathbf{\Delta}(x,n)$ are normally not statistically registered and need to be estimated by a graduation of the available age-specific rates. Assuming appropriate values

[6] This limitation has been noted by Ledent (1978, 1980).

of $\delta_{ij}(x)$ have been estimated[7] and are placed in matrix $\boldsymbol{\Delta}(x,n)$, the following formula calculates the more realistic probabilities of surviving[8]:

$$\mathbf{P}(x) = \left[\mathbf{I} + \left[n - \boldsymbol{\Delta}(x,n)\right] \circ \mathbf{M}(x,n)\right]^{-1} \left[\mathbf{I} - \Delta(x,n) \circ \mathbf{M}(x,n)\right] \tag{6}$$

The next step is to calculate life table survivors to exact age x in matrix $\mathbf{l}(x)$. Life table survivors to exact age $x + n$ are calculated as the product of the survivors to exact age x and the probability of surviving from exact age x to exact age $x+n$. Note, again, the order of the matrixes in this step. The survivors at exact age zero, or the radix of the life table, are arbitrary; here, they are set to 1.0, placed on the main diagonal of matrix $\boldsymbol{l}(x)$.

$$\begin{aligned} \mathbf{l}(x+n) &= \mathbf{P}(x)\mathbf{l}(x) \\ \boldsymbol{l}(0) &= diag(1.0) \end{aligned} \tag{7}$$

The total number of years lived by those surviving from exact age x to exact age $x+n$ during period T is the person-years of a life table in matrix $\mathbf{L}(x,n)$. It is important to recognize that the person-years lived are equivalent to the life table's stationary population. This equivalency is instrumental in formulating a practical apparatus for age-classified population projections, as will be shown.

An often-used approximation for calculating the person-years lived is to assume the decrements between exact age x and $x+n$ to be linear (all decrements occur exactly at midpoint). The approximation is shown in equation 8:

$$\mathbf{L}(x,n) \approx \frac{n}{2}\left[\mathbf{l}(x) + \mathbf{l}(x+n)\right] \tag{8}$$

However, the accurate formula for single and the multiple decrements uses the MDT correction factor $\delta_{ij}(x)$ again, as in equation 9:

$$\begin{aligned} \mathbf{L}(x,n) &= \mathbf{l}(x) \circ \boldsymbol{\Delta}(x,n) + \mathbf{l}(x+n) \circ \left[n - \boldsymbol{\Delta}(x,n)\right] \\ x &= 0,\ldots,(z-n). \end{aligned} \tag{9}$$

Again, the matrixes are multiplied elementwise (the Hadamard product).

A different approach is required for the last open-ended age group. Assuming again a stationary population for this age interval z, ∞, the formula for this open-ended age interval becomes

$$\mathbf{L}(z,\infty) = \left[\mathbf{M}(z,\infty)\right]^{-1} \mathbf{l}(z) \tag{10}$$

Adding $\mathbf{L}(x,n)$ from last to first age yields the matrix $\mathbf{T}(x)$, the total number of years lived beyond age x (equation 11):

$$\mathbf{T}(x) = \sum_{y=x}^{z} \mathbf{L}(y,n) \tag{11}$$

[7] For procedures to estimate the Mean Duration of Transfer, see Greville (1943), UNPD (1982, 31), Preston, Heuveline, and Guillot (2001, 44–47 and 82–84), and Schoen (2006, 18).

[8] The mathematical operator ∘ denotes the Hadamard product, or elementwise multiplication of two matrixes.

The multiregional life table in its basic from is completed by calculating the age- and region-specific life expectancies,[9] $\mathbf{V}(x)$, by dividing the total years lived beyond age *x* by the survivors to exact age *x*.

$$\mathbf{V}(x) = \mathbf{T}(x)\mathbf{l}(x)^{-1} \tag{12}$$

The additional dimension of demographic analysis, here, space or region, provides additional and valuable insights into the demographic system under study. As Willekens and Rogers (1978) show, lifetime can now be analyzed with two different perspectives, namely, life history by region of birth/origin and life history by region of residence/destination (see, for example, Willekens and Rogers (1978) and Rogers (1995)). As this paper is concerned with population projections and modeling international migration, this aspect of multiregional demography is not further discussed.

For the preparation of population projections by the cohort-component method, survivorship proportions are the most important element. They are the life-table equivalent of the factor or survivorship proportion needed to calculate the transition/survival of the population in age group *x*, *x+n* over an interval of *n* years forward *and* redistribute the population through migration spatially. Survivorship proportions are calculated as proportions of adjacent age groups of person-years lived. Three cases need to be distinguished: the survival of births born during the projection interval, closed age groups, and the last two age groups.

The survivorship proportion of the first age group determines the survival of births born during the interval of *n* years, forming the first age group (0, *n*) of the population at time *t+n* at the end of the period (equation 13). This special survivorship proportion is denoted with symbol *l* as the left superscript (as in *lower* Lexis triangle).

$${}^{l}\mathbf{S}(0) = \mathbf{L}(0)\left[n\mathbf{l}(0)\right]^{-1} \tag{13}$$

The survivorship proportion of the middle age groups is calculated as shown in equation 14.

$$\mathbf{S}(x,n) = \mathbf{L}(x+n,n)\left[\mathbf{L}(x,n)\right]^{-1} \tag{14}$$

The survivorship proportion of the last, open-ended age group includes two age groups of the population at time *t*: the last closed age group (*z-n*, *n*) and the last open-ended age group (*z*, ∞), shown in equation 15. Survivors of these two age groups become the new open-ended age group at time *t+n.*

$$\begin{aligned}\mathbf{S}(z-n,\infty) &= \mathbf{T}(z)\left[\mathbf{T}(z-n)\right]^{-1} \\ &= \mathbf{L}(z,\infty)\left[\mathbf{L}(z,\infty)+\mathbf{L}(z-n,n)\right]^{-1}\end{aligned} \tag{15}$$

The matrix representation of the survivorship proportions in a multiregional model is shown in equation 16.

[9] The usual symbol for life expectancy, e_x, is replaced in this paper by v_x and the corresponding matrix symbols.

$$\mathbf{S}(x,n) = \begin{bmatrix} s_{11}(x,n) & s_{21}(x,n) & \cdots & s_{m1}(x,n) \\ s_{12}(x,n) & s_{22}(x,n) & \cdots & s_{m2}(x,n) \\ \vdots & \vdots & \vdots & \vdots \\ s_{1m}(x,n) & s_{2m}(x,n) & \cdots & s_{mm}(x,n) \end{bmatrix} \tag{16}$$

A closer review of equation 16 reveals some interesting and useful properties. The survivorship proportions in a regional system are a manifestation of the underlying mortality in each region *and* the spatial/regional redistribution through migration. In effect, the survivorship proportions in matrix $\mathbf{S}(x,n)$ are the product of an underlying gross survivorship proportion $\hat{s}_i$ (introduced in equation 18) that is characteristic of a region[10] and proportions that govern the redistribution of survivors. We define the spatial redistribution proportion, τ_{ij}, as the proportion of the population in region *i* that will survive and redistribute to region *j*:

$$\tau_{ij} = \frac{K_{ij}}{K_i}; \sum_i^m \tau_{ij} = 1.0 \tag{17}$$

The (known) survivorship proportion, s_{ij}, can then be written as the product of the yet unknown gross survivorship proportion, $\hat{s}_i$, and the spatial redistribution proportion, τ_{ij}:

$$s_{ij} = \hat{s}_i \tau_{ij} \tag{18}$$

All the spatial/regional redistribution proportions for each region/column sum to unity, and the gross survivorship proportions must be identical for all elements in each column/region. It follows that the gross survivorship proportion is the sum of all the survivorship proportions of a region.

$$\begin{aligned} \sum_i^m s_{ij} &= \sum_i^m \hat{s}_i \tau_{ij} \\ &= \hat{s}_i \sum_i^m \tau_{ij} \\ \sum_i^m s_{ij} &= \hat{s}_i \end{aligned} \tag{19}$$

Define a matrix $\hat{\mathbf{S}}(x,n)$ containing the gross survivorship proportions with region-specific gross proportions on the main diagonal:

$$\hat{\mathbf{S}}(x,n) = \begin{bmatrix} \hat{s}_1(x,n) & 0 & \cdots & 0 \\ 0 & \hat{s}_2(x,n) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \hat{s}_m(x,n) \end{bmatrix} \tag{20}$$

[10] It follows from the Markovian nature of the survivorship proportions that the underlying mortality, and hence the survivorship proportion, is the same for all elements in a column of the $\mathbf{S}(x,n)$ matrix.

The gross survivorship proportions will later be utilized for calculating the deaths by age and region of origin and destination that are associated with the assumed age-specific mortality rates for each step of the multiregional projection exercise.

1.4. Multiregional Population Projections I

This section first summarizes the classic methodology of multiregional population projections. The section then suggests some useful extensions that can provide more details about the multiregional dynamics involved.

The cohort-component population projection method calculates the future population at time *t+n* in three steps. First, the population at time *t+n* is forward-projected by multiplying all but the last two age groups (the last-but-one and last, closed age groups) by the corresponding survivorship proportions. A second step is required to forward-project the last two age groups.[11] Finally, births during the projection interval are calculated separately (see section 2.2) and forward-survived to the end of the projection period, where they form the first age group (o, *n*) at time *t+n.*

Leslie (1945) introduced elegant and compact notation for a multiregional projection model that combines all three steps into one, and thus lends itself to a simple notational form of the projection process. However, for this project, it was decided to employ the more elaborate approach of multiregional projection with three steps (Willekens and Rogers 1978, 62). This approach affords more transparency, enables the explicit calculation of vital and migratory events, and avoids the large overhead of empty matrix cells in Leslie's approach.[12]

The main body of multiregional projections concerns the closed age groups of the base population, for example, all age groups except for the last two (age groups (*z-n*, *n*) and (*z*, ∞)). The population in these age groups at time *t+n* in vector $\mathbf{k}^t(x,n)$ is forward-projected by left-multiplying it with the corresponding survivorship proportions, $\mathbf{S}^T(x,n)$ (equation 21). The projected population at time *t+n* and age groups (*n, n*) through (*z-n, n*) is also in vector format.

$$\mathbf{k}^{t+n}(x+n,n) = \mathbf{S}^{t,n}(x,n)\ \mathbf{k}^t(x,n) \qquad x = 0,n,\ldots,z-2n \tag{21}$$

To arrive at the last, open-ended age group at the end the projection interval at time *t+n*, the two last age groups of the population at time *t+n* are involved,[13] that is, the last-but-one age group (which becomes part of the last age group at time *t+n*), plus the last, open-ended age group at time *t*. The projection is afforded by multiplying the sum of the last two population age groups by the survivorship proportion, $\mathbf{S}^T(z-n,\infty)$, calculated according to equation 15.

---

[11] This step in the projection process was treated incorrectly in earlier treatises on multiregional projection methodology, following a simplified assumption about survival beyond the last age group (Keyfitz 1968; Willekens and Rogers 1978; Rogers 1995).

[12] For example, in a single-sex population model with 21 age groups and three regions, about 93 percent of all the matrix elements are zero. The same model with two sexes would have about 97 percent of all elements set to zero.

[13] Early texts on multiregional demography (Willekens and Rogers 1978; Rogers 1995) as well as some ecological models have simplified the survival of the oldest age groups by assuming that the last, open-ended age groups will survive only *n* more years—an unrealistic assumption for most human populations.

$$\mathbf{k}^{t+n}(z,\infty) = \mathbf{S}(z-n,\infty)\left[\mathbf{k}^{t}(z-n,n) + \mathbf{k}^{t}(z,\infty)\right] \tag{22}$$

Finally, the projection of the first age group of the population at *t+n* requires the number of births during period T as input (equation 23). The births are forward-survived by applying the survivorship proportion of the lower Lexis triangle, ${}^{l}\mathbf{S}^{T}(0)$.

$$\mathbf{k}^{t+n}(0) = {}^{l}\mathbf{S}^{T}(0)\mathbf{b}^{T} \tag{23}$$

The methodology presented so far produces future populations by age and residence at the end of each projection interval. With a simple expansion of the population vector into a diagonal matrix, it is possible to also show where the projected population originated. This information may prove useful, inter alia, for creating and projecting migrant stock information in addition to population in each region. The expansion consists of placing the vector of population by age and region, $\mathbf{k}^{t}(x,n)$, into a diagonal matrix, $diag\left(\mathbf{k}^{t}(x,n)\right)$:

$$diag\left(\mathbf{k}^{t}(x,n)\right) = \begin{bmatrix} k_1^t(x,n) & 0 & \cdots & 0 \\ 0 & k_2^t(x,n) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & k_m^t(x,n) \end{bmatrix} \tag{24}$$

By repeating the projection steps described above but with the diagonal population matrix $diag\left(\mathbf{k}^{t}(x,n)\right)$ instead of the population vector, a population matrix $\mathbf{K}^{t+n}(x+n,n)$ at time *t+n* is obtained.

$$\mathbf{K}^{t+n}(x+n,n) = \mathbf{S}^{T}(x,n)\ \ diag\left(\mathbf{k}^{t}(x,n)\right) \tag{25}$$

The resulting population matrix $\mathbf{K}$ contains the population in region *i* at time *t+n* by region of origin *j* at time *t*.

. $$\mathbf{K}^{t+n}(x+n,n) = \begin{bmatrix} k_{11}^{t+n}(x+n,n) & k_{21}^{t+n}(x+n,n) & \cdots & k_{m1}^{t+n}(x+n,n) \\ k_{12}^{t+1}(x+n,n) & k_{22}^{t+n}(x+n,n) & \cdots & k_{m2}^{t+n}(x+n,n) \\ \vdots & \vdots & \vdots & \vdots \\ k_{1m}^{t+n}(x+n,n) & k_{2m}^{t+n}(x+n,n) & \cdots & k_{mm}^{t+n}(x+n,n) \end{bmatrix} \tag{26}$$

The main diagonal elements k_{ii} are those persons in region *i* who have survived from age group (*x*, *n*) to age group (*x+n*, *n*) and remained in the same region (called stayers). The off-diagonal elements in each row $k_{ij;\,j\neq i}$ are those people who have moved into region *i* from other regions and survived (called movers or immigrants). The sum of each row in matrix $\mathbf{K}$ is the total number of people who survived in region *j* (equation 27):

$$k_j^{t+n} = \sum\nolimits_i^m k_{ij}^{t+n} \tag{27}$$

The off-diagonal elements in each column, $k_{ij;i\neq j}$, are those people who have left region *i* and survived (called movers or emigrants). In other words, they are the surviving population of region *i* irrespective of their residence at the end of the projection period. The column sum k_i^{t+n} is then the total number of people who have survived to time point *t+n* and originated in region *i* (equation 28), but who are now in region *j*.

$$k_i^{t+n} = \sum\nolimits_j^m k_{ij}^{t+n} \tag{28}$$

1.5. Multiregional Population Projections II

The classic multiregional model adds movements/transitions into another region or state to the decrements into an absorbing state (mortality). Movements or transitions, as the case may be, enter the model as emigration intensities or rates, ensuring at the same time the proper calculation as exposure/occurrence rates. Such a formulation is transparent and consistent for any given observation period, as the emigration out of one region is, at the same time, an equivalent immigration into another region. But what happens if the regions involved exhibit differential demographic dynamics in the future? Then, for instance, the assumption of a constant number of migrants would mean decreasing emigration rates for sending countries with a growing population, and increasing emigration rates for sending countries with decreasing population size. Similarly, constant emigration rates would generate increasing absolute numbers of emigrants for countries with growing populations, and decreasing figures for countries with shrinking populations.

What about the countries that receive migrants? They are not explicitly treated in the standard multiregional model. They appear as passive receptors of migration flows originating elsewhere. This emigration bias[14] in the standard multiregional projection model is a serious impediment to a more realistic and flexible formulation of assumptions for international migration. The simple and practical alternative suggested in Buettner and Muenz (2017) will briefly be repeated here.

At least two challenges need to be addressed. The first is to include explicitly the dual nature of migration flows as being emigration for sending and immigration for receiving countries/regions. Such a perspective turns the standard multiregional projection methodology into a nonlinear model. Consequently, the projection requires iterations for its calculation. The second challenge is not mathematical, but conceptual. It seems problematic to formulate migration assumptions in terms of age-specific emigration rates. This is especially true for assumptions about immigration. We suggest that migration assumptions for population projections should use instead population-based indicators, such as total migration or crude migration rates, which are a better communication device and, most importantly, a better way to include immigration. In a separate step, the migration assumptions, formulated in absolute numbers or crude migration rates, still need to be translated into the corresponding age-specific rates needed by the underlying multiregional model.

[14] Demographers have pointed to this problem as early as when the method was initially developed (Feeney 1973). Similar criticism was later repeated and further elaborated and extended (Plane 1993; Courgeau 2006; Dion 2013; Bohnert et al. 2015). However, suggested solutions have mostly been formal so far, not practical, or restricted to subnational entities.

The translation of emigration rates into admission rates and vice versa is shown in the following.

The crude emigration and admission rates are defined, respectively, as:

Emigration (emission) rates	Admission rates	
$e_{ij} = \frac{E_{ij}}{\bar{K}_i}$	$a_{ij} = \frac{E_{ij}}{\bar{K}_j}$	(29)

A given emigration/admission rate is transformed into the corresponding admission/emigration rate:

Emigration to admission rate	Admission to emigration rate	
$a_{ij} = e_{ij} \cdot \frac{\bar{K}_i}{\bar{K}_j}$	$e_{ij} = a_{ij} \cdot \frac{\bar{K}_j}{\bar{K}_i}$	(30)

The relations in formulas 29 and 30 are valid for the base observation period. But if sending and receiving regions follow different demographic trajectories, the balance between the emigration and admission rates is no longer guaranteed. In Buettner and Muenz (2017), the harmonic mean between interacting regions is suggested as a simple balancing device (for assumptions formulated in absolute numbers or relative terms). Other approaches for balancing the emission and admission of migrants are of course possible; some can be found in Bijak (2011), Bohnert et al. (2015), Courgeau (2006), Dion (2013, 2017), Feeney (1973), Plane (1982), and Schoen (2006). Here, we present the harmonic mean updating mechanism for crude migration rates.

The consolidated crude migration rate at period *T* (h_{ij}^T) is calculated by multiplying a reference emigration rate—here, for the reference or base period of the projection—by the harmonic mean of population change between the base and projection periods, as shown in equations 31 and 32.

$$h_{ij}^T = e_{ij}^{ref} \frac{2 \cdot r_j^T}{\left(r_i^T + r_j^T\right)} \tag{31}$$

$$r_i^T = \frac{P_i^T}{P_i^{ref}} \tag{32}$$

At this step, we have obtained crude emigration rates from region *i* to region *j* at projection period T in the future that reflect the demographic dynamics (here, population size) of the sending (emigration) and receiving (admitting) countries by means of harmonic mean updating.

It has been suggested that the use of crude rates or total migration figures is preferred for the formulation of migration assumptions (Buettner and Muenz 2017). However, using crude rates or even total migration figures as a communication device must not compromise the formulation of the underlying projection model. The multiregional model employs age-specific emigration intensities as driving forces. Obviously, when one concept is used at the implementation level and another for formulating the projection assumptions, one needs to be translated into the other. For a mathematical explanation, see Buettner and Muenz (2017).

2 Events in a Multiregional Model

2.1. Demographic Events Gap

Standard cohort-component population projections for a single entity—country or region—are usually presented as a device to project an initial population by age (and sex) into the future. The direct outcomes of that exercise are future populations by age (and sex), plus births. If international migration is included in terms of net migration, this information is provided exogenously. When the results are presented, the size and age composition of future populations and synthetic summary indicators, such as life expectancy and total fertility, are discussed. If migration is specified in rates, the resulting absolute values are often omitted or only briefly mentioned.

The multiregional projection model has been developed with a similar limitation of the absolute demographic events that are associated with the projection process. The standard single-state and multiregional projection methodologies rarely, if at all, develop the mathematical apparatus for recovering the demographic events of death and migration. This section fills that gap.

2.2. Births

Births are a main factor of population renewal. However, as covered in textbooks on multiregional projections, the calculation of births is sometimes fraught with errors.[15] This paper provides a correct formula that follows standard practice in using an age-period format for age-specific fertility and population at risk calculation (for the single-decrement case, see Preston, Heuveline, and Guillot 2001, 124).

In the multiregional case, age- and region-specific fertility rates are first arranged on the main diagonal of the m-by-m matrix F(x). Multiplying the matrix of fertility rates by the average number of person-years in each age group gives the number of births by age of mother and region.

$$\mathbf{F}(x) = diag\left(b_i(x)\right) \tag{33}$$

$$\mathbf{b}^T(x,n) = \mathbf{F}^T(x,n)\bar{\mathbf{k}}^T(x,n) \tag{34}$$

At this point in the projection process, the population at risk, $\bar{\mathbf{k}}^T(x)$, for the current projection period may not be available, but can easily be calculated, as in equation 35.

$$\begin{aligned} \bar{\mathbf{k}}^T(x,n) &\approx \frac{n}{2}\left(\mathbf{k}^t(x) + \mathbf{k}^{t+n}(x)\right) \\ &\approx \frac{n}{2}\left(\mathbf{k}^t(x) + \mathbf{S}^T(x-n)\mathbf{k}^t(x+n)\right) \\ x &= \alpha, \ldots, \beta - n \end{aligned} \tag{35}$$

Summing births by age across the reproductive lifespan (usually from age α = 15 to age β = 50) yields a vector with the total number of births by region:

[15] Willekens and Rogers (1978) and Rogers (1995) incorrectly use an age-cohort format instead of the correct age-period format for specifying fertility rates and the population at risk.

$$\mathbf{b}^T = \sum_{x=\alpha}^{\beta-n} \mathbf{b}^T(x,n) \tag{36}$$

2.3. Deaths

The classic multiregional projection arithmetic projects, by region, people who survive in the same or other regions. In this schema, the actual amounts of movements between regions (migration or movement into another region) or deaths (movement into an absorbing state) are "hidden." The generating matrix of multiregional models, $\mathbf{M}(x,n)$, and the corresponding matrix of survivorship proportions, $\mathbf{S}(x,n)$, combine the effects of deaths and moving into another region. The elements on the main diagonal encompass the combined effects of moving-off (deaths) and moving-out (emigration) for each region (found in the columns of matrixes $\mathbf{M}(x,n)$ and $\mathbf{S}(x,n)$). Calculating deaths requires separating the two forces. As for the generating matrix of multiregional models, when adding the elements of each column, we obtain the underlying death rates of that column/region, without the effects of migratory movements (see section 1.2). The same operation on the $\mathbf{S}(x,n)$ matrix, for example, summing the elements of each column, results in gross survivorship proportions, $\hat{s}_i(x,n)$, that represent the underlying or gross survivorship proportions of that column/region. The individual gross survivorship proportions may be arranged in a row vector, $\hat{\mathbf{s}}(x,n)$, $\hat{\mathbf{s}}(x,n)$ or the diagonal matrix, $\hat{\mathbf{S}}(x,n)$, as shown in equations 20, 37, and 38.

$$\hat{\mathbf{s}}(x,n) = \left(\sum\nolimits_j^m s_{1j}(x,n) \ \cdots \ \sum\nolimits_j^m s_{mj}(x,n)\right) \tag{37}$$

$$\hat{\mathbf{S}}(x,n) = diag(\hat{\mathbf{s}}(x,n)) \tag{38}$$

Death, that is, the transition into an absorbing state from which return is not possible, is the logical complement of (gross) survival. Expressed formally, the proportion of people dying in region i in age group x during the period is

$$w_i(x,n) = 1 - \hat{s}_i(x,n) \tag{39}$$

For the whole multiregional model/system, the proportion dying may be arranged in a vector, $\mathbf{w}(x,n)$:

$$\mathbf{w}(x,n) = \left(1-\hat{s}_i(x,n) \ \cdots \ 1-\hat{s}_m(x,n)\right) \tag{40}$$

Deaths are now calculated as the product of the population in region i at the beginning of the period and the proportion dying during the period (and in region i at the beginning):

$$d_i(x,n) = k_i(x,n)\, w_i(x,n) \tag{41}$$

In matrix notation, equation 41 can be written as:

$$\mathbf{d}(x,n) = \mathbf{k}(x,n)\mathbf{w}(x,n) \tag{42}$$

For the calculation of deaths that are associated with a given set of survivorship proportions, we define the matrix $\mathbf{W}(x,n)$ as the complement of the matrix of gross survivorship proportions:

$$\mathbf{W}(x,n) = 1 - \widehat{\mathbf{S}}(x,n) \tag{43}$$

The deaths in a multiregional system are then obtained as:

$$\begin{aligned} \mathbf{d}^T(x,n) &= \left[1 - \widehat{\mathbf{S}}^T(x,n)\right]\mathbf{k}^t(x,n) \\ &= \mathbf{W}^T(x,n)\,\mathbf{k}^t(x,n) \end{aligned} \tag{44}$$

Because of the Markovian nature of the multiregional projection model, deaths calculated this way refer to all deaths in the population that was in region *i* at the beginning of the period, and the mortality of the region of origin also operates for all outmigrants.

Deaths calculated from a cohort-component projection model are, as the migrants (see section 2.4), available in a period-cohort format and need to be transformed to the common period-age format, if required. In principle, this transformation splits the cohort deaths into lower and upper Lexis triangles and then re-assigns them. It is often deemed sufficient to split the events in each age group in half. More elaborate and accurate procedures for the proper redistribution of events between different observation formats (age-cohort and age-period) are discussed in Wilmoth et al. (2007).

2.4. Migrants

The number of migrants is not a direct factor in a multiregional projection model—neither as input nor as output. Information about migratory movements is in the form of age-specific emigration rates, which are required to combine the effects of mortality and migratory movement. The direct outcome of a projection step is the population by age and region—only the births have been calculated (directly, as in this paper, or indirectly, as in most multiregional models and especially in the Leslie format). It is possible to retrieve the "hidden" component of death from the multiregional model (section 2.3). The procedures for obtaining the numbers of migrants by age and region are outlined in this section. But why is it useful to do the extra steps and calculate this component of demographic change explicitly?

At least two reasons can be brought forward. First, it is simply more convenient to formulate migration scenarios in terms of absolute numbers or crude rates. It is also the easiest to communicate.[16] And to implement the scenarios, the numbers of migrants by age and region need to be available for every projection step. Second, "recovering" migrants by age and region enables insights into the impact of demographic change—namely, population aging—on the age composition of migrants.

The number of migrants can be calculated in two slightly different ways:

- The number of migrants that are alive at the end of the interval in a region different from where they left. This number of migrants is relatively simple to calculate and will be developed further.
- The total number of migrants during the interval that includes those who died during the interval. This number of migrants requires additional steps to calculate, such as to account for the competing risks

[16] Although it is common to specify assumptions for mortality and fertility in population projections in terms of synthetic indicators, such as life expectancy at birth and total fertility (measures that refer to an average individual), this is not the case for migration. The corresponding measure summarizing the migration experience of an average person—the gross migraproduction rate—has not found any significant traction in textbooks or practical applications.

between dying and migrating; it will be developed in a forthcoming publication. Such a separate consideration of migrants included in the number of deaths is not standard fare, even in the classical cohort-component projections using net migration.

$$\mathbf{O}(x,n) = \begin{bmatrix} 0 & s_{21}(x,n) & \cdots & s_{m1}(x,n) \\ s_{12}(x,n) & 0 & \cdots & s_{m2}(x,n) \\ \vdots & \vdots & \vdots & \vdots \\ s_{1m}(x,n) & s_{2m}(x,n) & \cdots & 0 \end{bmatrix} \tag{45}$$

For the calculation of migrants in a multiregional system, the off-diagonal elements of the matrix of survivorship proportions $\mathbf{S}(x,n)$ represent the proportion of the population at the beginning that survives to the end of the period and is in another region, as arranged in matrix $\mathbf{O}(x,n)$. This implies that the emigration proportions in matrix $\mathbf{O}(x,n)$ produce only surviving migrants. Migrants who entered the destination region and then died before the end of the period are still assigned to their region of origin as deaths. Emigrants are calculated by left-multiplying the outmigration proportions by the population matrix.

$$\mathbf{E}(x,n) = \mathbf{O}(x,n)\mathbf{K}(x,n) \tag{46}$$

Migrants (emigrants) calculated from a cohort-component projection model, as for deaths, are available in a period-cohort format and need to be transformed to the common period-age format, if required. In principle, a similar approach as for deaths may be used for migration data.

3. Accounting for Nonlinear Decrements

3.1. From Rates to Probabilities

Population projections are firmly based on the well-known demographic tool, the life table. A crucial element of life table construction is the translation of observed demographic events or rates into corresponding probabilities as the basis of the life table.

Three assumptions about the distribution of events have been widely used; two assume an underlying distribution of events, and third a local fit:

- Uniform distribution
- Exponential distribution
- Approximation by a local fit.

The uniform assumption is straightforward, as it assumes that all events of decrement (for example, deaths and outmigration) are equally distributed through each age interval. This assumption may therefore be used for simple exploratory analysis and relatively narrow age groups. It produces erroneous results, however, for larger age groups (more than one year) and high rates of decrement.

The exponential assumption posits that the forces of decrement are constant through each age interval. This assumption fits rapidly increasing decrements well (mortality at higher ages, for instance), but underestimates the dynamics of decrement in cases of declining forces (Preston, Heuveline, and Guillot 2001, 46).

The local option is based on a local function that is assumed to represent the distribution of the forces of decrement locally by fitting this through neighboring age groups. For the software used for this paper, an estimation variant was chosen that goes back to Greville (1943), was theoretically underpinned by Chiang (1960), and used in United Nations Population Division (1982, 2013). The chosen approach is shown for age-specific mortality rates, but can be extended to migration as well.

3.2. First Age Group

The translation of age-specific decrement rates into probabilities for the first age group poses a special challenge due to the extreme compression of mortality (probably less pronounced for migration) into the first days and months of life. Under this circumstance, the above-mentioned approaches do not work. Consequently, there are several approaches to accounting for the extreme nonlinearity of forces of decrement in this age group (Preston, Heuveline, and Guillot 2001, 47–48.). One approach is to borrow the information from a set of reference life tables, by regressing the correction factor against the known age-specific mortality rates. Preston, Heuveline, and Guillot (2001, table 3.3) present regression results for the first two age groups of abridged life tables, namely the age groups (0,1) and (1,4).

For this project, we repeat that approach, but for the age group (0,5), which is the common age format for (global) population projections. Table 2 shows the parameters for calculating ${}_5\delta_0$ based on the ${}_5m_0$ from the updated Coale-Demeny model life table in UNPD (2015a) for Coale-Demeny CD West.

Table 2: Function Parameters for Estimating the Mean Duration of Transfer for Age Group (0,5)

Sex	Function to estimate ${}_5a_0$	No.
Males	${}_5a_0 = \begin{cases} -0.7427\,{}_5\mathrm{m}_0 + 0.9966 & {}_5m_0 \geq 0.06 \\ 4768.1*{}_5\mathrm{m}_0{}^3 - 735.36*{}_5\mathrm{m}_0{}^2 + 38.319*{}_5\mathrm{m}_0 + 0.2703 & {}_5m_0 < 0.06 \end{cases}$	(47)
Females	${}_5a_0 = \begin{cases} -0.8145*{}_5\mathrm{m}_0 + 1.1002 & {}_5m_0 \geq 0.06 \\ 4825.5*{}_5\mathrm{m}_0{}^3 - 768.63*{}_5\mathrm{m}_0{}^2 + 40.828*{}_5\mathrm{m}_0 + 0.3264 & {}_5m_0 < 0.06 \end{cases}$	(48)

For cases where population projections are performed for one sex only (males and females combined), a simple averaging of male and female estimates may be sufficient.

3.3. Middle Age Groups

For the middle age groups, the Greville approach is summarized in formula 47, substituting the usual symbol, ${}_na_x$, with a symbol representing the Median Duration Transfer concept, ${}_n\delta_x$. The ${}_nm_x$ originally referred to the age-specific mortality rate at *x* to *x+n*, but in the multiregional model, it stands for all decrements more generally.

$$
\begin{aligned}
{}_n\delta_x &= \frac{n}{2} - \frac{n^2}{12}\left({}_n m_x - {}_n k_x\right) \\
{}_n k_x &= \frac{1}{2n} Ln\left(\frac{{}_n m_{x+n}}{{}_n m_{x-n}}\right)
\end{aligned}
\tag{49}
$$

The factor *k* corrects for the deviation of the actual force of mortality from the uniform distribution assumption; that is, *k* accounts for decelerating/accelerating changes in the forces of decrement. This formula cannot be used for the first and last closed age groups (the last-but-one age group) and the last, open-ended age group. However, observing that ${}_n k_x$ changes relatively slowly at old ages, the ${}_n a_x$ for the last closed age group may be estimated, without much loss of accuracy, by "borrowing" the ${}_n k_x$ from the predecessor age group.[17]

$$
{}_n k_{z-n} \approx {}_n k_{z-2n} \tag{50}
$$

This shortcut approach can be improved by extrapolating the age-specific values of *k* to the last closed age group. Preliminary tests suggest that a quadratic function is promising, but a short exponential extension might be more appropriate.

3.4. Last, Open-Ended Age Group

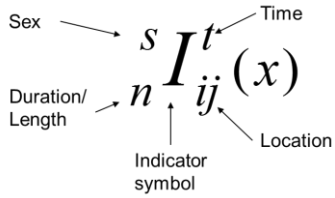
For the last, open-ended age group, the average number of years lived by those dying during that open-ended age interval, ${}_\infty a_z$, is the same as the remaining life expectancy, as everybody eventually will die beyond age *z*. The common approximation is then (in terms of single-decrement life tables):

$$
{}_\infty\delta_z = {}_\infty L_z = T_z = e_z \approx \frac{l_z}{{}_\infty m_z} \tag{51}
$$

[17] In MORTPAK's life table program, the last closed age group is estimated using the age-specific mortality rate for that age group and the *k* of the preceding age group.

Annex A. Notation and Symbols

A.1. Notation



Total demographic stock variables, like population and events such as births, deaths, or migration, are written in uppercase letters. Crude demographic rates are written in corresponding lowercase letters; age-specific rates have age indexes added in parentheses.

Multiregional demographic models make heavy use of compact matrix and vector notation.[18] Uppercase bold-face letters denote matrixes; their elements are written as lowercase letters with subscripts. Vectors are denoted by lowercase bold-face letters, and their elements by lowercase letters with subscripts.

Matrix multiplication is commonly denoted by writing the matrix symbols in sequence, for example, **AB**, where **A** and **B** are two matrixes. It is important to note that matrix multiplication is not commutative and hence the order of the matrixes matters (**AB≠BA**).

Some operations of multiregional algebra include scalar factor matrixes. In these cases, multiplication is done elementwise, known as the Hadamard product and written **A**∘**B**.

The extension of demographic notation to include an additional regional dimension made it necessary to make a distinction between the notation used in single-decrement life tables and that used in multistate life tables and models. For life tables, (exact) age is denoted by a right subscript, and the width of age groups or the length of time periods by a subscript to the left of the symbol. For the multiregional case, age is identified by adding the (exact) age in parentheses.

The regional dimensions are generally noted by subscripts *i* and *j*, where *i* indicates the region of origin, and *j* stands for the region of destination. To reduce the notational complexity, time or period is indicated only where necessary, by a right superscript.

The following table assembles all the symbols used in this document.

[18] Matrix notation also supports the implementation of such models in matrix-oriented programming languages, like R.

A.2. List of Symbols

Symbol	Description
K_i	Total population in region i
$K_i(x,n)$	Population in region i in age group (x, n)
$\overline{K}_i^T$	Total person-years lived by the population in region i during period T $(t, t+n)$
$\overline{K}_i^T(x,n)$	Person-years lived by the population in region i during period T in age group (x, n)
B_i	Total births in region i
$B_i(x,n)$	Births in region i and age group x (x, n)
b_i	Crude birth rate in region i
$b_i(x,n)$	Age-specific birth rate in region i and age group (x, n)[a]
D_i	Total deaths in region i
$D_i(x,n)$	Deaths in region i in age group (x, n)
d_i	Crude death rate of region i
$d_i(x,n)$	Age-specific death rate of region i and age group (x, n)
E_{ij}	Total emigrants moving from region i to region j
$E_{ij}(x,n)$	Emigrants moving from region i to region j in age group (x, n)
e_{ij}	Crude emigration rate from region i to region j
$e_{ij}(x,n)$	Age-specific emigration rate from region i to region j and age group (x, n)
$e_i(x,n)$	Age-specific emigration rate from region i to all region j, age group (x, n)
$\delta_{ij}(x)$	Mean duration at transfer in age group (x, n) and moving from region i to j
$\boldsymbol{\delta}(x,n)$	Vector of mean duration at transfer in age group (x, n) for those remaining in region i. $\boldsymbol{\delta} = (\delta_{11}, \delta_{22}, \cdots, \delta_{mm})$
$\boldsymbol{\Delta}(x,n)$	Matrix of Mean Duration at Transfers in age group x by region of origin i and region of residence j
$\mathbf{M}(x,n)$	Generating matrix with age-specific mortality and emigration rates for age group (x, n)
$\mathbf{I}$	Identity matrix; m-by-m matrix with 1.0 at the main diagonal, zero otherwise
$p_{ij}(x)$	Age-specific probability of surviving from exact age x to $x+n$ and from region i to region j
$\mathbf{P}(x)$	Matrix with age-specific probabilities of surviving from exact age x to $x+n$
$l_{ij}(x)$	Survivors to exact age x in region of residence j born in region i
$\mathbf{l}(x)$	Matrix of survivors to exact age x (x, $x+n$) in an $m \times m$ regional system
$L_{ij}(x,n)$	Person-years lived in age group (x, n) and region j by those born in region i

List of Symbols (continued)

Symbol	Description
$\mathbf{L}(x,n)$	Matrix of person-years lived in age group (x, n) in an m*m regional system
$T_{ij}(x)$	Total number of years lived beyond exact age x by region of birth i and region of residence j
$\mathbf{T}(x)$	Matrix with total number of years lived beyond age x
$\mathbf{V}(x)$	Matrix with life expectancies at exact age x
$s_{ij}(x,n)$	Age-specific survivorship proportion for age group *(x, n)* from region i to region j
$\mathbf{S}(x,n)$	Matrix with age-specific survivorship proportions for age group *(x, n)*
${}^{l}\mathbf{S}(0)$	Survivorship proportion from birth (exact age 0) to the end of period T (survivorship in the first lower Lexis triangle)
$\mathbf{b}(x,n)$	Births to residents in region i in age group *(x, n)*[b]
$\mathbf{E}(x,n)$	Matrix with age-specific emigration rates from region i to region j
$\mathbf{F}(x)$	Diagonal matrix of age-specific fertility rates of mothers in age group *(x, n)*
A_{ij}	Total admissions of migrants from region i into region j
$A_{ij}(x)$	Admissions of migrants from region i into region j in age group x
a_{ij}	Crude admission rate of migrants from region i to region j (immigration rate)
$a_{ij}(\mathrm{x})$	Age-specific admission rate of migrants from region i to region j (immigration rate)
$\boldsymbol{A}(x)$	Matrix with age-specific admission rates
$\mathbf{k}^{t}(x,n)$	Vector of population by region and age group *(x, n)*
$\overline{\mathbf{k}}^{T}(x,n)$	Vector with person-years lived by regions for population at age group *(x, n)* during period T (average population)
$diag\left(\mathbf{k}^{t}(x,n)\right)$	Diagonal population matrix with elements of vector $\mathbf{k}^{t}(x)$ in the main diagonal
$\mathbf{K}^{t}(x)$	Matrix with population by destination and origin in age group *(x, n)* at time t
$\mathbf{O}(x,n)$	Matrix with outmigration proportions (emigration proportions)
$\mathbf{W}(x,n)$	Matrix with proportions dying for age group *(x, n)*
$\widehat{s}_{i}(x,n)$	Gross survivorship proportions for age group *(x, n)*
$\widehat{\mathbf{s}}(x,n)$	Vector with gross survivorship proportions for age group *(x, n)*
$\widehat{\mathbf{S}}(x,n)$	Matrix with gross survivorship proportions for age group *(x, n)*
$ms_{i}(x)$	Proportionate age-specific migration rate (migration schedule)
gmr_{i}	Gross migraproduction rate for region i

List of Symbols (continued)

Symbol	Description
	Indexes
i	Region of origin; $i = 1, 2 \ldots, m$
j	Region of destination; $j = 1, 2, \ldots, m$
m	Number of regions
n	Width of an age group or length of an observation/projection period
x	Exact age
x, n	Age group from age x to $x+n$ (years)
α	First age of reproduction (usually set to $\alpha = 15$)
β	Last age of reproduction (usually set to $\beta = 50$)
t	Date, point in time (usually year)
T	Period from t to $t+n$
z	Last age; age at the beginning of the last, open-ended age group

a. Most demographic textbooks use the symbol $f(x)$. For consistency with the notation used in this paper, the symbol $b(x,n)$ replaces it.

b. In Willekens and Rogers (1978) and Rogers (1995), the symbol ***B*** is used to denote the matrix of age-specific fertility rates discounted by the proportion of live births surviving to the end of the period. As this paper is committed to also calculating the vital events in a multiregional model, the old meaning (which "hides" births) was replaced with the vital events (without discounting for survival).

Bibliography

Abel, G. J. 2009. *International Migration Flow Table Estimation*. University of Southampton. http://eprints.soton.ac.uk/69577/1.hasCoversheetVersion/Thesis.pdf.

———. 2013. "Estimating Global Migration Flow Tables Using Place of Birth Data." *Demographic Research* 28 (March): 505–46. https://doi.org/10.4054/DemRes.2013.28.18.

Abel, G. J., and N. Sanders. 2014a. "Quantifying Global International Migration Flows." *Science* 343 (6178): 1520–22. https://doi.org/10.1126/science.1248676.

———. 2014b. "Quantifying Global International Migration Flows. Supplementary Materials: Database S1 and S2." *Science* 343 (6178): 1520–22. doi:10.1126/science.1248676. http://science.sciencemag.org/content/343/6178/1520/suppl/DC1.

Bijak, Jakub. 2011. *Forecasting International Migration in Europe: A Bayesian View*. Vol. 24. The Springer Series on Demographic Methods and Population Analysis. Dordrecht: Springer Netherlands. doi:10.1007/978-90-481-8897-0.

Bohnert, N., J. Chagnon, S. Coulombe, P. Dion, and L. Martel. 2015. *Population Projections for Canada (2013 to 2063), Provinces and Territories (2013 to 2038): Technical Report on Methodology and Assumptions*. National Population Projections Team, Statistics Canada, Ottawa, Ontario.

Buettner, T., and R. Muenz. 2016. "Comparative Analysis of International Migration in Population Projections." KNOMAD Working Paper 10, World Bank, Washington, DC.

———. 2017. "Modeling Alternative Projections of International Migration." KNOMAD Working Paper 30, World Bank, Washington, DC.

Caswell, H. 2001. *Matrix Population Models: Construction, Analysis, and Interpretation*, 2nd ed. Sunderland, MA: Sinauer Associates.

Chiang, Chin Long. 1960. "A Stochastic Study of the Life Table and Its Applications: I. Probability Distributions of the Biometric Functions." *Biometrics* 16 (4): 618. doi:10.2307/2527766.

Courgeau, Daniel. 2006. "Mobility and Spatial Heterogeneity." In *Demography: Analysis and Synthesis,* vol. 1, edited by Graziella Caselli, Daniel Courgeau, Jacques Vallin, and Guillaume J. Wunsch, 287–88. Academic Press.

Dion, P. 2013. "An Alternative Projection Model for Interprovincial Migration in Canada." No. WP 14.3, Joint Eurostat/UNECE Work Session on Demographic Projections, Rome, Italy. http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.11/2013/WP_14.3_01.pdf.

Dion, P. 2017. "An Alternative to Fixed Transition Probabilities for the Projection of Interprovincial Migration in Canada." *Population Research and Policy Review*, June. doi:10.1007/s11113-017-9440-6.

Feeney, Griffith. 1973. "Two Models for Multiregional Population Dynamics." *Environment and Planning A*. doi:10.1068/a050031.

Greville, T. N. E. 1943. "Short Methods of Constructing Abridged Life Tables." *Record of the American Institute of Actuaries* XXXII (17): 29–42.

Hill, K. 1990. *PROJ3S - A Computer Program for Population Projections: Diskettes and Reference Guide*. Washington, DC: World Bank.

Keyfitz, Nathan. 1968. *Introduction to the Mathematics of Population*. Addison-Wesley.

Ledent, Jacques. 1978. “Some Methodological and Empirical Considerations in the Construction of Increment-Decrement Life Tables.” IIASA Research Memorandum RM-78-025, International Institute for Applied Systems Analysis, Laxenburg, Austria.

———. 1980. “An Improved Methodology for Constructing Increment-Decrement Life Tables from the Transition Perspectives.” IIASA Working Paper WP-80-104, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Leslie, P. H. 1945. “On the Use of Matrices in Certain Population Mathematics.” *Biometrika* 33 (3): 183–212. https://doi.org/10.1093/biomet/33.3.183.

Plane, D. A. 1982. “An Information Theoretic Approach to the Estimation of Migration Flows.” *Journal of Regional Science* 22 (4): 441–56. doi:10.1111/j.1467-9787.1982.tb00769.x.

Plane, D. A. 1993. “Requiem for the Fixed-Transition-Probability Migrant.” *Geographical Analysis* 25 (3): 211–23.

Preston, Samuel H., Patrick Heuveline, and Michel Guillot. 2001. *Demography: Measuring and Modeling Population Processes*. Malden, MA: Blackwell Publishers.

R Development Core Team. 2016. *R: A Language and Environment for Statistical Computing*. Vienna, Austria. http://www.r-project.org.

Rogers, Andrei. 1995. *Multiregional Demography: Principles, Methods and Extensions*. Chichester, NY: Wiley.

Rogers, A., and F. Willekens. 1976. “Spatial Population Dynamics.” In *Papers of the Regional Science Association*, vol. 36, 1–34. Budapest, Hungary.

Schoen, R. 1979. “Calculating Increment-Decrement Life Tables by Estimating Mean Durations at Transfer from Observed Rates.” *Mathematical Biosciences* 47 (3–4): 255–69. https://doi.org/10.1016/0025-5564(79)90041-5.

———. 1988. *Modeling Multigroup Populations*. Springer Series on Demographic Methods and Population Analysis. Boston, MA: Springer US. doi:10.1007/978-1-4899-2055-3. http://link.springer.com/10.1007/978-1-4899-2055-3.

———. 2006. *Dynamic Population Models*, vol. 17. Springer Series on Demographic Methods and Population Analysis. Dordrecht: Springer Netherlands. doi:10.1007/1-4020-5230-8. http://link.springer.com/10.1007/1-4020-5230-8.

UNPD (United Nations Population Division). 1982. “Model Life Tables for Developing Countries.” *Population Studies* (77).

———. 2013. “Mortpak for Windows (Version 4.3).” New York. http://www.un.org/en/development/desa/population/publications/mortality/mortpak.shtml.

———. 2015a. *Extended Model Life Tables*, version 1.3. New York. https://esa.un.org/unpd/wpp/Download/Other/MLT/.

———. 2015b. *Trends in International Migrant Stock: The 2015 Revision*. POP/DB/MIG/Stock/Rev.2015. New York.

Willekens, Frans. 2008. “Models of Migration: Observations and Judgments.” In *International Migration in Europe: Data, Models and Estimates*, edited by James Raymer and Frans Willekens, 117–47. Chichester, England: John Wiley & Sons, Ltd.

Willekens, F., and A. Rogers. 1978. *Spatial Population Analysis: Methods and Computer Programs*. *Research Report RR-78-18*. International lnstitute for Applied Systems Analysis, Laxenburg, Austria. http://pure.iiasa.ac.at/825.

Wilmoth, John R, K Andreev, D A Jdanov, and D A Glei. 2007. “Methods Protocol for the Human Mortality Database (Version 5).” University of California Berkeley, CA; Max Planck Institute for Demographic Research, Rostock. http://www.mortality.org/.

KNOMAD
Global Knowledge Partnership on Migration and Development