Fertestr: an R package for fertility estimation

José H C Monteiro da Silva[1]
Everton E C de Lima[2]
Helena C Castanheira[3]
Patrick Gerland[4]

Abstract
Several countries still rely on indirect methods for fertility estimation because their vital statistical systems are either incomplete or non-existent. Therefore, their fertility estimates are mostly drawn from surveys and censuses or from incomplete vital statistics, which often require additional adjustment using indirect demographic methods. The goal of the *fertestr* package is to compile the most commonly used indirect methods for data quality assessment and fertility estimation in a single R package. In this paper, we present several functions implemented and examples of their applications at both the national level, and for population subgroups and subnational level.

Introduction
Reliable fertility estimates are the key to deriving demographic scenarios and projections. In an ideal scenario, age-specific fertility rates can be directly estimated using birth counts by the mother's age and female population exposure by age. However, several countries still lag in the collection of good vital statistics information, or they have birth numbers that are incomplete and/or under-reported (Mikkelsen et al. 2015). Therefore, these countries still rely on censuses, surveys, or indirect methods to estimate overall fertility levels and trends, and by age groups and other socio-demographic characteristics of the mothers or households.

Most of the indirect methods used are documented in the *Tools for demographic estimation (Moultrie et al. 2013)*, a successor of the UN Manual X, which also has excel spreadsheets for replication of the methods. Additionally, to our knowledge, there is no compilation of these methods using a single R package. Hence, our main goal is to develop an R package with the most commonly used indirect fertility estimation methods that can be applied to data from low- and middle-income countries. Given the increase in the number of R users among demographers and population scientists, this resource would accelerate the process of data analysis and fertility estimation, including its application at the subnational level.

In this paper, our main goal is to present the package and some of its functions so that future users can become familiar with it and use it appropriately.

Methods

[1] Population Studies Center, University of Pennsylvania, Philadelphia, PA, USA – corresponding author: josehcms@sas.upenn.edu
[2] Population Studies Center and College of Philosophy and Human Sciences, University of Campinas, Campinas, Brazil
[3] Latin American and Caribbean Demographic Centre, Population Division of the United Nations Economic Commission for Latin America and the Caribbean, Santiago, Chile
[4] United Nations Population Division, United Nations, New York, USA

The package includes a broad range of fertility estimation methods, among which we list the four main methods below. Beyond these listed ones, the package also includes methods for evaluating the plausibility of children ever born reported by mothers, and the El Badry method for assessing the relationship between missing data and zeros in the reporting of children ever born.

Brass PF

The Brass PF series method is based on the approach that if fertility has been constant, one can compare the cohort (children ever born, parity) fertility with period fertility (children had over the last 12 months, for example) (Brass 1964). In the package, we include two versions of the Brass PF method: the original one (Brass 1964) and the updated in the Manual X that includes the Colae-Trussel multipliers (United Nations 1983).

Relational Gompertz PF

The relational Gompertz PF method implemented in the package represents an updated version of Brass PF that adjusts the level of fertility schedules collected on censuses and surveys (using the question on recent births) by using parity information. The implementation follows the steps described in the *Tools for Demographic Estimation* (Moultrie et al. 2013).

Reverse Survival

The idea of the reverse survival method implemented in the package is to estimate the number of births and total fertility rates up to 15 years prior to censuses and surveys from the population counts (by age) of children (ages 0-14) and women (ages 15-64). The rationale is to apply a set of survival functions to reverse survive these children to their year of birth, and the women to each of the previous 15 years. The implementation also closely follows the one proposed in the *Tools for Demographic Estimation*, but we also implemented an additional approach that constructs five-year estimates instead of generating yearly estimates constructs 5-year estimates. We proposed this version because, in many situations, we only have abridged life tables, so we do not need to apply any graduation or model method to obtain single-year life tables to perform their reverse survival.

Own-children method

The own-children fertility method is currently under implementation, and its rationale is similar to that of the reverse survival method. In addition, it can provide age-specific fertility rates. However, for this method, we need an extra piece of information: the number of children listed by each mother's age group. The method implemented in the package follows the procedure used in the Manual X (United Nations 1983).

Summary and next steps

The number of R users has been growing among the demographic community and national statistics offices because of its advantage in making data analysis more efficient and less time consuming. Therefore, the goal of the *fertestr* package is to equip users with a broad set of tools for fertility estimation, offering them the opportunity to test different approaches and select the one that best suits their needs. The integration of these tools into the R ecosystem further promotes reproducibility and collaboration within the demographic community.

More information about the package, and the documentation of the functions, and its codes can be found on the package's GitHub page: https://github.com/josehcms/fertestr.

The next steps of our work include the revision of the functions and implementation of the *cohort parity comparison with vital registration data*, also described in the *Tools for demographic estimation.*

References

Brass, W. (1964). Uses of census or survey data for the estimation of vital rates. Presented at the African Seminar on Vital Statistics, Addis Ababa: Economic Commission for Africa.

Mikkelsen, L., Phillips, D. E., AbouZahr, C., Setel, P. W., De Savigny, D., Lozano, R., & Lopez, A. D. (2015). A global assessment of civil registration and vital statistics systems: monitoring data quality and progress. *The Lancet*, *386*(10001), 1395–1406. https://doi.org/10.1016/S0140-6736(15)60171-4

Moultrie, T., Dorrington, R., Hill, A., Hill, K., Timæus, I., & Zaba, B. (2013). *Tools for Demographic Estimation*. Paris: International Union for the Scientific Study of Population (IUSSP).

United Nations. (1983). *Manual X: Indirect Techniques for Demographic Estimation*. New York: United Nations, Department of Economic and Social Affairs.