DEMOGRAPHIC MODELS

Philippe Wanner
Forum suisse pour l'Étude des migrations, Université de Neuchâtel, Neuchâtel, Switzerland

Josianne Duchêne
Institut de démographie, Université catholique de Louvain, Louvain-la-Neuve Belgium

Keywords: modeling, mortality models, nuptiality models, models of conception, models of birth, fertility curves, intermediate factors, migration models, migration curves, population growth, demography, stable population, demographic projections, mathematic functions.

Contents

Summary

A model is a simplified and most often mathematical representation of reality, frequently used in demography to understand mortality, nuptiality, fertility, migration and demographic evolution. We can distinguish the fitting of demographic curves and the modeling of processes in order to understand, explain and describe the behavior of populations. In this paper, different kinds of models are discussed.

For instance, mortality laws were for a long time very simplistic and not always accurate. On the contrary, some of the most recent models are complex, but they do not always give more information. In fact, there is no mathematical function that is able to translate exactly current mortality patterns.

Concerning fertility, fitting curves allows, for example, the forecasting of fertility trends. It is also useful to improve the measure of fertility indicators when the quality of data is not perfect. Other models suppose that general fertility rate is a function of fecundity, nuptiality, contraception use and effectiveness and induced abortion or are oriented toward the analysis of fertility control.

Less attention has been paid to migration modeling, perhaps because of the absence of underlying biologically governed processes. In return, migration models have to take into account the spatial dimension, and also socio-economic factors. They are therefore generally rich and multidisciplinary. Contrary to mortality or fertility, there are few models of age-specific migration curve adjustment.

Finally, models of unconstrained population growth tend to be based on exponential curves depending only on the initial population size, time, and the intrinsic rate of growth. Lotka has established the basic ideas of the stable population theory.

1. Modeling and demography

Modeling is one of the possible forms of scientific approach, often used in social science and in particular in demography. A *model* is a simplified and most often mathematical representation of reality, which includes the variables that may account for the explanation of the aspects taken into account.

This simplification process leads to an improvement in the perception of reality and allows operations that can not be conducted otherwise, for example the simulation by the manipulation of a limited number of parameters, in order to measure their impact on the studied dimension. Indeed, one of the main characteristics of a model is that it must be easier to handle than the reality. Modeling is then useful in formulating and solving some complex problems.

A condition for modeling is to express the reality and its dimensions in precise terms. The expressions or terms used to describe an observed reality have often ambiguous meanings. This does not allow the scientist to understand the underlying mechanisms and to measure the studied dimensions in an accurate way. In the mathematical language, symbols are strictly and uniquely defined and they answer to precise and explicit rules of manipulation. Hypotheses must also be explicitly and precisely formulated. If it is not the case, modeling is not valid.

The earliest models were *physical*. If they were usually fairly easy to construct, they also were difficult to modify. After the physical models, the second kind of models were the graphical ones, easier to construct but more abstract and then less understandable. More recently, *symbolic* models came into use. This type of models is easier to manipulate compared to physical ones. They are abstract and the number of variables that may be included is unlimited. Symbolic models have had different stages of development. The first ones were empirical; then in the 1850s came curve-fitting models. Thereafter they became deterministic, and finally stochastic. Demographic models described in this paper are most often symbolic ones.

Demography is particularly suitable for the use of models, and among them numerical ones because the studied events – such as births, deaths, marriage or divorce - and their intensity – expressed by rates or proportions – are numerical. However, despite the large number of models, numerous demographic problems are still to be studied. Modeling has still a long way to go.

1.1 The definition and the limits of modeling

Modeling must be distinguished from *quantification* and from the *use of statistical techniques* to measure dimensions. A confusion between the three concepts often occurs, because all of them use mathematical language. Quantification is used to measure a phenomenon, while statistical techniques allow the validation of hypotheses. These two tools are often used to validate a model.

A confusion is also often made between *model* and *theory*. A model refers to the description of a phenomenon or a category of phenomena. It implies a smaller and less general construction than a theory, which is sometimes less rigorous, but lays claim to a greater explicative value.

The more a model fits correctly the reality, the more we can admit that the mechanisms it describes are correct. When the model correctly predicts several of different observable characteristics and when it can be used in different conditions, then model and theory are merging.

1.2 The different kind of demographic models

In what follows, only specific demographic models are presented. Multidisciplinary models, or models developed in other discipline but also used in demography are not taken into account. We will also distinguish between models concerning the demographic evolution as a whole (stable population models, for example) and models concerning one or more phenomena. Models that deal with the population evolution start from an initial population showing some defined characteristics, and simulate, according to beforehand established rules, the events that modify the size and structure of the population.

Models may vary according to different classifications:

- The way the model takes time into account. When the probabilities of occurrence of an event fluctuates in function of time, the model is considered as a *dynamic* one. If that is not the case, the model is considered as *static*.
- The dimension studied. A *macro-model* deals with groups of individuals presenting the same characteristics, while a *micro-model* deals with each individual considered separately.
- We can also make the distinction between *deterministic* and *stochastic* models. A deterministic model assumes that, if each individual of a group of N persons has a probability p to experiment the event E, this event will occur exactly pN times. A stochastic model assumes that the number of times the event E will occur may differ from one simulation to another and from the expected number pN, but follows a random distribution with a mean equal to pN.

The choice of the model depends on the aim of modeling, as well as the means at our disposal: a deterministic macro-model gives an estimation free from sampling error, but needs the computation of a large range of information. A stochastic micro-model needs less information, is more flexible, but is limited due to the sampling error. There is, in fact, no static micro-model.

Fitting curves are another type of model, which are used when the identification of predictive processes is considered as more important than the precise analysis of explicative mechanisms. This is the case, in particular, when the main target of the modeling is forecasting.

Finally, it must be stated that models presented below are among the most famous demographic models. But other models also exist, and may be useful for some purposes. The target of each scientific approach is to find the best tool to solve the problems under study. Therefore, modeling and models have to be used with care.

2. Mortality models

Mortality is the demographic phenomenon that has been the most frequently studied. The first mortality law was suggested in the third Century by Ulpien, in order to compute pension rates and life annuities and to implement the Lex Falcidia. This mortality law supposed a constant decrease of the survivals numbers. More recently, in 1662, John Graunt produced a survival table based both on observed data for the youngest and on a geometrical progression between 6 and 76 years. Graunt's model was built at the end of the large epidemics periods. At that time, age was not considered as a factor of mortality risk and Graunt's model was only dependent of time. Some years after, demographers moved from time-dependent models to age-dependent ones. Today, all mortality models are age-dependent ones.

Mortality laws were for a long time very simplistic and not always accurate. On the contrary, some of the most recent models are complex, but they do not always give more information. In fact, there is no mathematical function that is able to translate exactly current mortality patterns.

Fitting curves are not the only type of mortality models. Some deterministic models were also developed, such as Hamer's model (1906) which dealt with the risk of measles mortality, or spatial models, such as Bartlett's model (1957) concerning the diffusion of epidemic waves. These models are not really considered as demographic ones, and are not discussed here.

2.1 The shape of the risk and survival curves.

The shape of the mortality curves is irregular and therefore models are almost complex (Figure 1). The *risk intensity* decreases during the first 10 years of life, before it increases in different ways according to the mortality level and the characteristics of the population. Thus, in almost all industrialized countries, one can observe between ages 20 and 30 a period of excess mortality, due to violent deaths. Moreover, in several countries, especially African ones, aids epidemics have led to a modification of mortality risk patterns. Finally, in the oldest-old ages (after 90) the increase rate of mortality risk is lower than expected.

The mathematical relationship between the risk function and other mortality functions is obvious, so the choice of the function is not important. In practice, fitting models are expressed in terms of mortality risk, survival probability or other derived functions (death distribution, instant risk, for example).

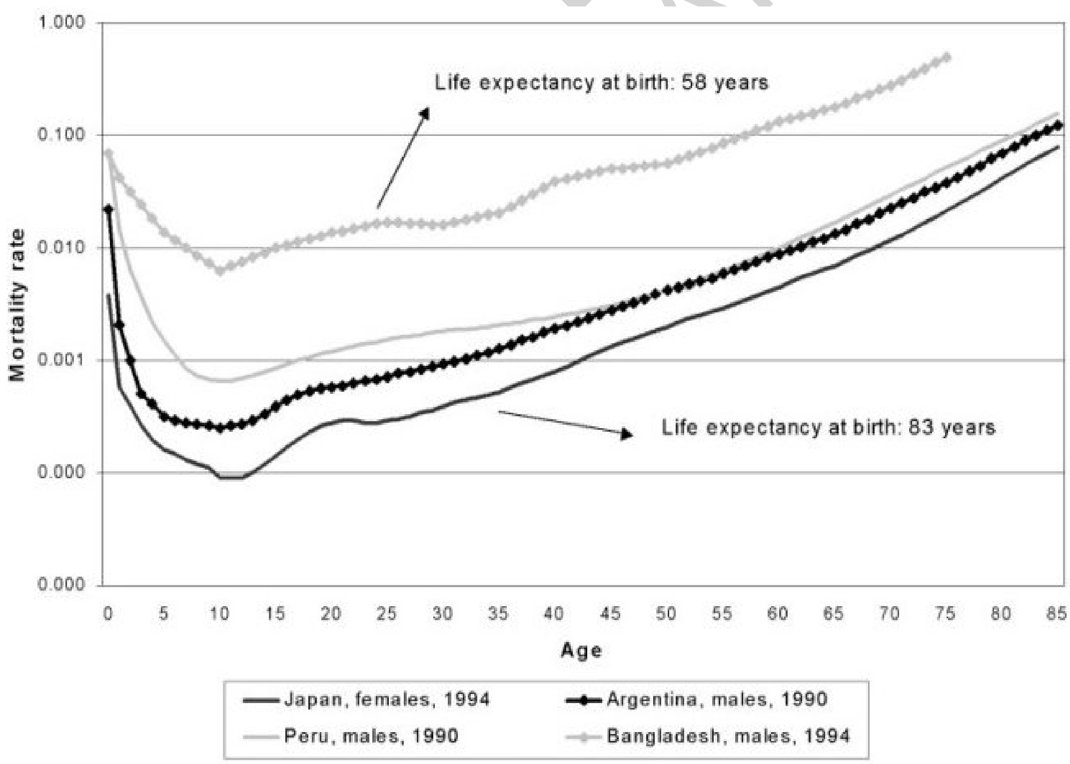

Figure 1: Some patterns of mortality risk, according to different levels of life expectancy

2.2 The first models of mortality

In 1662, John Graunt was one of the first to propose a distribution of deaths by age based on scientific assumptions. His work was built on registrations of deaths furnished by the London parishes. The survival function, underlying this distribution, is a geometrical progression between the ages of 6 and 76. From a root of 100 "quick conceptions", Graunt had estimated that only 64 persons would still be alive at the age of 6. After this age, the decrease of the number of survivors is equal to 1/3 (in other terms the proportion of deaths is 2/3) for each ten-years age group. This probability of mortality leads in fact to one survivor remaining at the age of 76.

At the same time (1670), the Dutch Jacob Van Dael proposed a nearly similar model, which assumed a division by two of the number of survivors in each 12 years age groups. In both cases, life expectancy at birth was less than 20. Both models also assumed that mortality risk is not increasing in function of age, in other terms age is not influencing mortality risk. The prevailing epidemics during the period before the formulation of these models is determinant to understand this time-dependent approach.

In 1671 De Wit was the first to express a mortality model for which life expectancy may vary. He also was the first to propose an age-dependent model. De Wit observed an increase of the mortality risk with age. His model is slightly more sophisticated: it assumes that the mortality risk between 63 and 73 is twice the risk between 3 and 53, as well as the risk between 73 and 80 is three times the one between 3 and 53.

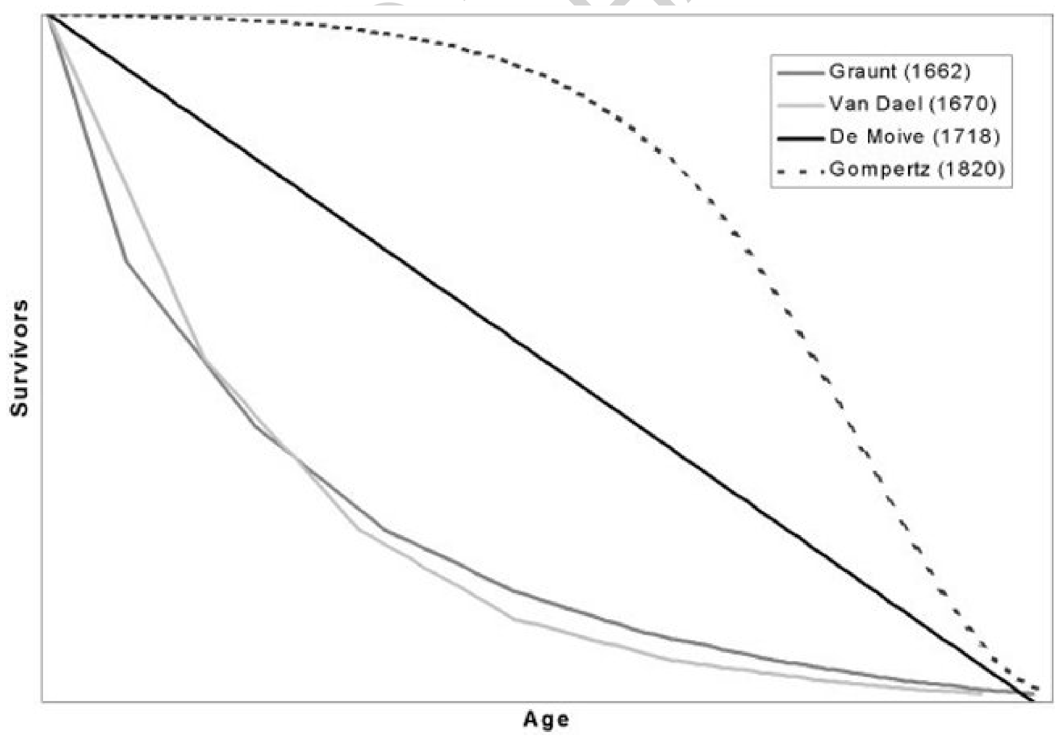

Figure 2: The form of the survival function, according to the first models of mortality

The list of such descriptive models dating from the 18th or 19th century, built in most cases on parish data, is quite long. One of these models, the well-known Abraham de Moive's model, is still in use. The French mathematician formulated, in 1718, a mortality law for which the number of deaths in a life table is the same for each age group. In this way, the survival function decreases at a linear pace (Figure 2), while the mortality risk expressed by the ratio between the number of deaths and the number of survivals, increases more and more quickly with age. De Moive was also the first to introduce the concept of limit of life span.

The Gompertz model is probably the best known and the most often used of all demographic models. Gompertz observed that a geometrical progression fits well the number of survivors when the increase in age follows an arithmetical progression (Figure 2). Moreover, Gompertz's function takes into account the deterioration of the physiological resistance, due to the increase in age. Gompertz's survivors function follows therefore the following expression: $l_x = kg^{q^x}$ where k, g and q are constants to be estimated.

William Makeham, in the 18th century, tried to improve this function and proposed in particular to add a multiplier term h^x. This modification represents the theory of partial forces of mortality, which suggests that during the oldest ages, the increase of the mortality risk is even faster than expected in the original function. Gompertz or Makeham's functions are often used to fit mortality curves, especially at the end of mortality tables, when the small size of surviving population requires fitting techniques. The Gompertz's function is also used for modeling in gerontology, biology and ecology.

2.3 Additive components mortality models

Quickly, demographers understood that an unique multiplicative or exponential function is not sufficient to fit the mortality curve. Models became more sophisticated and most often composed of two or more additive parts (See Appendice 1). The Danish actuary Thiele formulated, at the end of the 19th Century, a model that was covering all ages with three components. The first one deals with child mortality, the second with mid-age mortality, and the third with mortality at oldest ages. This approach is not far from some of the models proposed nowadays. For example, Siler's model (1979) is composed of three components and five parameters. The first component concerns child mortality, the second the hazard risk of death during life and the third (a Gompertz function) is an age-dependent component. Siler was inspired by his studies on animal populations and that is why his model ignores excess mortality due to violent deaths in the age range 20-30 (Figure 3).

One year after Siler, Heligman and Pollard (1980) demonstrated that a three components and eight parameters model may fit, with a high degree of accuracy, the mortality risk curve. The first component represents the child mortality, the second one the young adults mortality and the third the increase of mortality in older ages. Each parameter has a precise meaning: for example, one represents the intensity of infant mortality, another the mean age at violent death after 20. This model, often used for forecasting mortality, does not always give better results than Siler's one, which does not take into account violent deaths and deals with 3 parameters less, but fits the best old-age mortality (Figure 3).

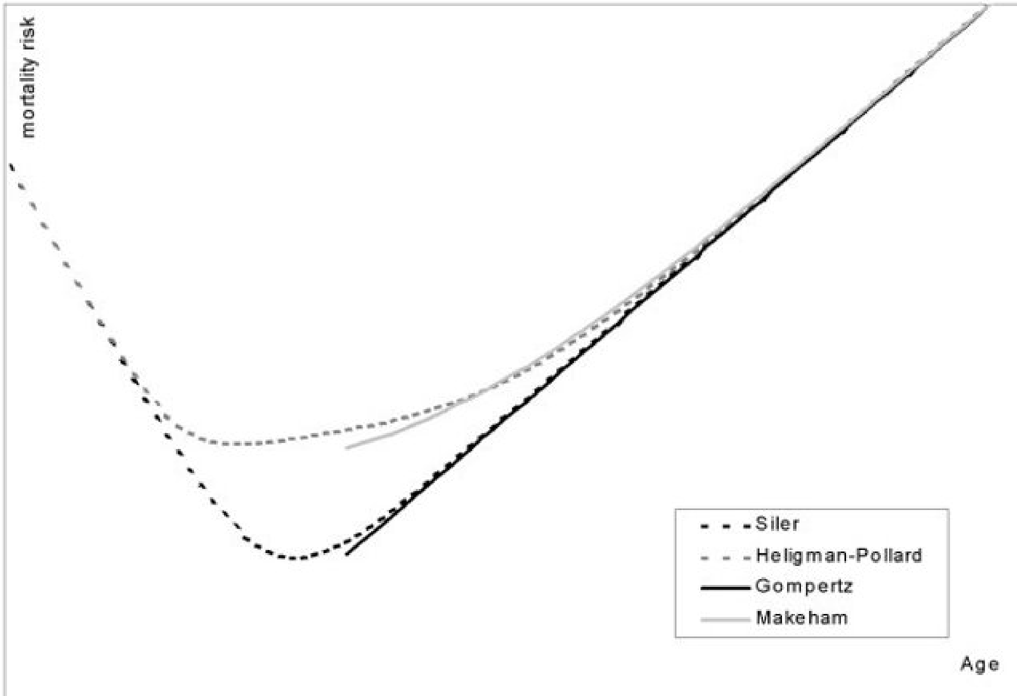

Figure 3: Adjustments of mortality risk, according to the most often used models

The comparison between models introduces the question about the aims of modeling. Improving the quality of a model often requires increasing the number of parameters. However, a model has to stay a simplification of reality and thus has to be quite simple. Moreover, the choice of a model depends on the purpose of modeling. For this reason, Siler's model can be adopted when the estimation of older-age populations should be accurate, while Heligman-Pollard's model is more useful for estimations of mortality at middle-age.

-
-
-

Bibliography

Bogue, D.J., Arriaga, E.E., Anderton, D.L. and Rumsey, G.W (1993). *Readings in population research* *methodology*. New York, New York, USA : Center: United Nations Population Fund [UNFPA], Chicago, Illinois : Social Development.

Coale A.J. (1972). *The Growth and Structure of Human Populations. A Mathematical Investigation*, 227 pp. Princeton, New Jersey, USA: Princeton University Press.

Henry L. (transl. by E. van de Walle) (1976). *Population: analysis and models*, 301 pp. London: Arnold.

Keyfitz N. (1968). *Introduction to the Mathematics of Population*, 450 pp. Reading, MA, USA : Addison-Wesley.

Manton K.G. and Stallard E. (1984). *Recent Trends in Mortality Analysis*, 342 pp. Orlando, Florida, USA: Academic Press.

Pollard J.H. (1973). *Mathematical Models for the Growth of Human Populations*, 186 pp. Cambridge, Great Britain: Cambridge University Press.

Rogers A (1995). *Multiregional demography: principles, methods and extensions,* 236 pp. Chichester, England : John Wiley and Sons.

Schoen R. (1988).*Modeling Multigroup Populations*, 308 pp. New York: Plenum Press.

Stillwell J. and Congdon P. (1991). *Migration models: macro and micro approaches.* 329 pp. London, England, Belhaven Press.

Biographical Sketches

Wanner Philippe, PhD in Demography, Researcher at the Swiss Forum for migration Studies at the University of Neuchatel. Research fields include demographic modeling and projection, mortality and fertility analysis. Author of different researches concerning Swiss population and mortality evolution.

Duchêne Josianne , PhD in Demography. Professor of Demography at the Institute of Demography. Université Catholique de Louvain, Louvain-la-Neuve. Author of researches concerning methodology, households and families, adult mortality and morbidity.