Autoregressive Integrated Moving Average (ARIMA) Models for Birth Forecasting

JOÃO LUIZ MAURITY SABOIA*

Autoregressive integrated moving average (ARIMA) models are developed for birth time series, and their relationship with classical models for population growth is investigated. Parsimonious versions for the ARIMA models are obtained which retain the most important pieces of information including the length of generation of the population. The technique is applied to birth time series data for Norway.

KEY WORDS: Birth time series; ARIMA models; Norwegian births; Population models; Forecasting.

1. INTRODUCTION

This article is intended both to improve present methods of forecasting births and to provide an alternative method of forecasting for countries with less available data. Although the time-series techniques used here (Box and Jenkins 1970) have been known for some time and have been applied to the study of populations in different ways (Hiorns 1972; Lee 1974; Pollard 1970; and Saboia 1974), our models are new in many aspects. As far as we know, this is the first attempt to justify the use of autoregressive integrated moving average (ARIMA) models for birth time series.

The point of departure is the classical renewal equation for births. Unlike the classical demographic models, we allow the vital rates to vary with time, and migration is permitted, as with real human populations. An ARIMA model is then developed. This model, however, has too many parameters and for practical reasons requires simplification. We show how such simplification can be accomplished while still obtaining a useful model. The simplified (parsimonious) model retains the most important pieces of information, including the length of generation of the population. In terms of classical theory, this corresponds to using the largest eigenvalues of the projection matrix while disregarding the smaller ones. Once the simpler ARIMA model is derived, forecasts, together with confidence intervals, can be obtained. The technique is then applied with success to birth time series for Norway.

Section 2 presents a very brief summary of both deterministic (Section 2.1) and probabilistic (Section 2.2) population models. In Section 3.1, an ARIMA model for the birth time series is introduced, and a simplified (parsimonious) form of this model is developed in Section 3.2. The technique is applied to real population data in Section 3.3. Finally, some comments as well as suggestions for future research are given in Section 4.

2. DETERMINISTIC AND STOCHASTIC DISCRETE MODELS FOR POPULATION GROWTH

2.1 Deterministic Discrete Models

In this section, we summarize some results of the deterministic discrete model for population growth and we reference the main results which are necessary to understand the development of the parsimonious model of Section 3.2. Most of this theory was developed by Bernadelli (1941), Lewis (1942), Leslie (1945, 1948a, 1948b), and more recently by Pollard (1966) and Goodman (1967).

Let us consider a female population at discrete points in time, say, $t = 0, 1, 2, \ldots$, and age-intervals corresponding to the unit intervals of time, say, $0, 1, 2, \ldots, M$. The usual practice is to consider only the female population in the prereproductive and reproductive age-intervals, say, $0, 1, 2, \ldots, N$ $(N < M)$.

Let $r(j)$ be the number of females born during period of time t, per female in the jth age-interval $(j = 0, 1, \ldots, N)$ at time t, who will be alive in the 0th age-interval at time $t + 1$. Let $s(j)$ be the proportion of females in the age-interval j at time t, who will be alive in the age-interval $(j + 1)$ at time $t + 1$. As is the usual practice, let us assume that the rates $r(j)$ and $s(j)$, $j = 0, 1, \ldots, N$, vary with age j but are independent of time t. Let $x_t(j)$ be the number of females alive in the jth age-interval at time t, and let $\mathbf{X}_t$ denote the column vector $\mathbf{X}_t = (x_t(0), x_t(1), \ldots, x_t(N))'$. Then the following relations hold

$$\mathbf{X}_t = \mathbf{A}\mathbf{X}_{t-1} , \tag{2.1}$$

or equivalently,

$$\mathbf{X}_t = \mathbf{A}^t\mathbf{X}_0 , \tag{2.2}$$

* João Luiz Maurity Saboia is Associate Professor, COPPE/Federal University of Rio de Janeiro, Caixa Postal 1191-ZC00-20000, Rio de Janeiro, RJ, Brazil. The research on which this article is based was conducted while the author was a research assistant at the Operations Research Center and a Ph.D. candidate at the Department of Industrial Engineering and Operations Research, University of California, Berkeley. It has been partially supported by the Office of Naval Research under Contract N00014-A-0200-1070, CAPES, and COPPE/Federal University of Rio de Janeiro, Brazil, with the University of California. The author wishes to express his gratitude to Professor Richard E. Barlow for the help given during this research. Thanks are also extended to Dr. Ronald Lee for helpful comments on an early version of the manuscript and to two referees and an associate editor for the suggestions.

© Journal of the American Statistical Association
June 1977, Volume 72, Number 358
Applications Section

where $\mathbf{A} = \{a_{ij}\}$ is the $(N+1) \times (N+1)$ matrix

$$\mathbf{A} = \begin{bmatrix} r(0) & r(1) & \dots & r(N-1) & r(N) \\ s(0) & 0 & \dots & 0 & 0 \\ 0 & s(1) & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & s(N-1) & 0 \end{bmatrix} . \quad (2.3)$$

The eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_{N+1}$ of $\mathbf{A}$ can be determined by solving the following equation (Goodman 1967):

$$\sum_{j=0}^{N} \phi(j)\lambda^{-(j+1)} = 1 , \quad (2.4)$$

where

$$\phi(j) = r(j)p(j) \quad (2.5)$$

and

$$\begin{aligned} p(j) &= 1 \quad \text{for} \quad j = 0 , \\ &= \prod_{i=0}^{j-1} s(i) \quad \text{otherwise} . \end{aligned} \quad (2.6)$$

The three largest eigenvalues of $\mathbf{A}$ in absolute value are very important. The largest one, λ_1, can be interpreted as the eventual rate of increase of the female population. The next two largest eigenvalues in absolute value, λ_2 and λ_3, are related to the length of generation and, for real populations, are complex conjugates. If we write

$$\lambda_2 = be^{i\alpha} , \quad (2.7a)$$

$$\lambda_3 = be^{-i\alpha} , \quad (2.7b)$$

then the length of generation, L, can be approximated (Keyfitz 1972) by

$$L = 2\pi/\alpha . \quad (2.8)$$

If we now consider only females in the 0th age-interval, $x_t(0)$, which is approximately equal to the number of female births, then the renewal equation for births holds (Goodman 1967), i.e.,

$$b_t = \sum_{j=0}^{N} \phi(j)b_{t-j-1} , \quad (2.9)$$

where b_t is the number of females born during period of time t. For (2.9) to hold, $r(j)$, $j = 0, 1, \dots, N$, should now be understood as the number of females born during period t, per female in the jth age-interval at time t.[1] Equation (2.9) is the starting point for the discussion of Section 3.

2.2 Stochastic Discrete Models

The development of stochastic discrete models for population growth started recently and has followed two different lines. The first one supposes the vital rates to be invariant probabilities at the individual level. This procedure makes use of branching process results (Harris 1963), and the main results obtained are asymptotic first and second moments for the various age groups considered. (See Pollard (1966), (1969); Goodman (1968); and Schweder (1971).) The main problem with such an approach is the assumption of fixed probabilities. This implies prediction variances which are too small, so that Schweder (1971, p. 448), in the concluding remarks of his paper, observes that

> ... the source of projection deviation must rest mainly on year-to-year variation in death probabilities and birth distributions and in the error made when estimating these quantities. The pure randomness of population dynamics is of minor importance.

The second approach considers such yearly variations and treats vital rates as variables indexed on time. (See Sykes (1969); and Le Bras (1971).) Unfortunately, this approach has not yet been able to make a practical contribution to forecasting and developing confidence intervals for real populations. According to Lee (1947, p. 608), one reason for such failure is that

> [this approach] concentrates on the demographically interesting but quantitatively unimportant problem of internal correlations of the population projection matrix, and ignores the autocovariance structure of fertility variation as a process in time ... In the absence of reliable causal models, to ignore the autocovariance structure of fertility is to throw out the most useful available information.

Lee (1974) developed a linear model relating variations in fertility to variations in births, but, instead of supposing the fertility process to be serially uncorrelated as Sykes did, he assumed an autoregressive structure. To apply his model to real human populations raises a few difficulties. First, it makes use of fertility data to forecast the birth sequence. Although this procedure is closer to demographic theory and common practice than using births directly, it can constitute a problem when applying the model to populations with less available data. This is the case with most less developed countries, for which even data for birth time series are very difficult to obtain. Second, there is the problem of identifying an appropriate model for the fertility process (Lee discusses the difficulties of identifying a model for U.S. fertility (1974, Sec. 8)). Third, as we will show later, his model is not parsimonious. The concept of a parsimonious model which consists of using the smallest number of parameters for adequate representation was introduced by Tukey (1961).

Although fertility is very important by itself, the next sections will show how to ignore its autocovariance structure and analyze its influence on the birth sequence by studying directly the autocorrelation function of the birth time series. Also, we will show how to obtain a parsimonious model.

3. AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA) MODELS FOR THE BIRTH TIME SERIES

3.1 Nonparsimonious Models

In this section, we show how an ARIMA model for the female-birth time series $\{b_t\}$, $t = 1, 2, \dots, T$, can be obtained. As in a real human population, fertility and survival rates are allowed to vary with time. To make the

[1] This definition is slightly different from the previous one.

analysis simpler we suppose that the population is closed to migration but later relax this assumption.

The renewal equation for births is given by

$$b_t = \sum_{j=0}^{N} \phi_t(j) b_{t-j-1} , \tag{3.1}$$

where $\phi_t(j)$ is the *net maternity function* for the jth age-interval from birth to the tth time period,[2] and N is the largest reproductive age.

Let $\phi(j)$ be an average value for $\phi_t(j)$ during the period of time, $t = 1, 2, \ldots, T$. Thus we can write

$$\phi_t(j) = \phi(j) + e_t(j) \ , \quad j = 0, 1, \ldots, N \ , \tag{3.2}$$

where $e_t(j)$ is the change in the net maternity function for age j during period of time t from its average value $\phi(j)$. Thus we can rewrite (3.1) as

$$b_t = \sum_{j=0}^{N} \phi(j) b_{t-j-1} + \sum_{j=0}^{N} e_t(j) b_{t-j-1} \tag{3.3}$$

or, equivalently,

$$b_t = \sum_{j=0}^{N} \phi(j) b_{t-j-1} + n_t \ , \tag{3.4}$$

where

$$n_t = \sum_{j=0}^{N} e_t(j) b_{t-j-1} \ ; \tag{3.5}$$

i.e., the female-birth time series $\{b_t\}$ can be expressed as an autoregressive model in which the autoregressive coefficients are average values of the net maternity function during the period of time considered, and the noise time series $\{n_t\}$ corresponds to the number of births due to changes in the net maternity function.

Analysis of (3.5) leads us to expect the time series $\{n_t\}$ to be serially correlated. This is due not only to the fact that neighboring values n_t and n_{t+1} have common terms b_{t-j-1}, $j = 0, 1, \ldots, N - 1$, but also, as mentioned in Lee (1974), to the fact that the fertility process should be considered serially correlated. A class of linear models which incorporates the autocorrelation structure is the ARIMA models class of Box and Jenkins (1970).

Let us thus suppose that the disturbance time series $\{n_t\}$ can be represented by a stationary and invertible ARIMA (p_1, d_1, q_1) process, i.e.,

$$\Psi_1(B) n_t = \Theta_1(B) a_t \ , \tag{3.6}$$

where B is the backward shift operator such that $Bb_t = b_{t-1}$, $B^j b_t = b_{t-j}$; $\{a_t\}$, $t = 1, 2, \ldots$, is a sequence of independently distributed random variables having mean zero and variance σ_a^2 (white noise);

$$\Psi_1(B) = (1 - \phi_1 B - \ldots - \phi_{p_1} B^{p_1})(1 - B)^{d_1} = \Phi_1(B)\nabla^{d_1}$$

is the autoregressive polynomial;

$$\Theta_1(B) = 1 - \theta_1 B - \ldots - \theta_{q_1} B^{q_1}$$

is the moving average polynomial; and the roots of the equations $\Phi_1(B) = 0$ and $\Theta_1(B) = 0$ lie outside the unit circle (stationarity and invertibility conditions, respectively).

Multiplying both sides of (3.4) by $\Psi_1(B)$, we obtain

$$\Phi(B)\Phi_1(B)(1 - B)^{d_1} b_t = \Theta_1(B) a_t \ , \tag{3.7}$$

where

$$\Phi(B) = 1 - \sum_{j=0}^{N} \phi(j) B^{j+1} \ ; \tag{3.8}$$

i.e., the female-birth time series b_t can be expressed as an ARIMA model.

The main difficulty with (3.7) is that a very large number of parameters have to be estimated. (For human populations N is approximately 50.) In the next section we show how to solve this problem and attain *parsimony*, i.e., how to use the smallest number of parameters and still obtain a useful model.

Two final observations should be made here. First, since in practice the *sex ratio* does not vary considerably for different periods of time (being usually very close to the average value 1.05), we can multiply both sides of (3.4) by unity plus the average sex ratio and obtain an ARIMA model for the total birth time series, i.e., males and females considered together. Second, we can include migration in the model by simply adding a disturbance term to (3.1) representing births due to net migration of mothers.

3.2 Parsimonious Models

We start by considering the stationary and invertible ARIMA (p, d, q) model:

$$\Psi(B) X_t = \Theta(B) a_t \ , \tag{3.9}$$

where $\Psi(B) = (1 - \phi_1 B - \ldots - \phi_p B^p)(1 - B)^d = \Phi(B)\nabla^d$; $\Theta(B) = 1 - \theta_1 B - \ldots - \theta_q B^q$; $\{X_t\}$ is the time series being analyzed; and $\{a_t\}$ is a sequence of independent random variables having mean zero and variance σ_a^2.

Let

$$x_t = (1 - B)^d X_t \ , \tag{3.10}$$

where $\{x_t\}$ is a stationary time series.

Thus (3.9) can be written as

$$\Phi(B) x_t = \Theta(B) a_t \ . \tag{3.11}$$

It is a well-known result (Box and Jenkins 1970) that the *autocorrelation function* $\{\rho_k\}$ of the process $\{x_t\}$ satisfies the following relation:

$$\Phi(B) \rho_k = 0 \ , \quad k > q \ . \tag{3.12}$$

The solution to (3.12) depends on the roots of the equation $\Phi(B) = 0$. If we suppose that the roots are distinct (as in the case of real populations), the solution

[2] Notice that the definition of the net maternity function, $\phi_t(j)$, as given by (3.1), is slightly different from the conventional one. It takes into consideration the survival rates from birth to age j for someone born during time interval $t - j - 1$, as well as the fertility rate during period of time t for age j. The conventional definition considers survival rates during time-interval t.

to (3.12) is given by:

$$\rho_k = A_1G_1{}^k + A_2G_2{}^k + \ldots + A_pG_p{}^k \ , \qquad k > q - p \ , \quad (3.13)$$

where $G_1{}^{-1}, G_2{}^{-1}, \ldots, G_p{}^{-1}$ are the roots of $\Phi(B) = 0$, and $A_1, A_2, \ldots, A_p$ are constants which depend on the coefficients of the polynomial $\Theta(B)$. Each pair of complex conjugate roots $G_i{}^{-1}$, $G_j{}^{-1}$ can be considered together, in which case their combined contribution to (3.13) will be

$$Ab^k \sin (k\alpha + \beta) \ , \qquad (3.14)$$

where b is the absolute value and α the phase of the pair G_i, G_j; A and β are constants.

Since the ARIMA model considered is stationary (i.e., the roots of $\Phi(B) = 0$ lie outside the unit circle), we can conclude from these results that the autocorrelation function of the mixed model (3.11) consists of $q - p + 1$ initial values $\rho_0, \rho_1, \ldots, \rho_{q-p}$, followed by a mixture of damped exponentials and/or damped sine waves, whose nature is dictated by the roots of the equation $\Phi(B) = 0$ and the starting values. (Notice that if $p > q$, the whole autocorrelation function consists simply of a mixture of damped exponentials and/or damped sine waves.)

With these considerations in mind, let us return to the linear model of (3.7). We first notice that the polynomial $\Phi(B)$ of (3.8) is identical to the polynomial in λ of (2.4), provided that we substitute B by λ^{-1}, so that the roots of $\Phi(B) = 0$ are the inverse of the roots of (2.7), i.e., the roots of $\Phi(B) = 0$ are the inverse of the eigenvalues of the projection matrix $\mathbf{A}$ of section 2.1. This observation is the key point for obtaining a parsimonious ARIMA model for the female time series $\{b_t\}$. As mentioned in Section 2.1, the positive eigenvalue λ_1 is the largest in absolute value among all eigenvalues, and for real human populations it is very close to one. The next two largest eigenvalues in absolute value, λ_2 and λ_3, also play a very important role. For real human populations, λ_2 and λ_3 are complex conjugates, smaller than one in absolute value, and can be used to determine the length of generation (see (2.8)). To illustrate the importance of the eigenvalues λ_1, λ_2, and λ_3, an approximate solution to the deterministic renewal equation for the births can be obtained by making use of only these three eigenvalues (Keyfitz 1972).

Based on this discussion, a first approximation to the ARIMA model of (3.7) is given by

$$\Phi_1(B)\Phi_2(B)\Phi_3(B)(1 - B)^{d_1+1}b_t = \Theta_1(B)a_t \ , \quad (3.15)$$

where $\Phi_1(B)$ is the autoregressive polynomial of the disturbance $\{n_t\}$; $\Phi_2(B)$ is such that the roots of $\Phi_2(B) = 0$ are $\lambda_2{}^{-1}$ and $\lambda_3{}^{-1}$, respectively; $\Phi_3(B)$ is such that the roots of $\Phi_3(B) = 0$ are $\lambda_4{}^{-1}, \lambda_5{}^{-1}, \ldots, \lambda_{N+1}{}^{-1}$; and the root $\lambda_1{}^{-1}$ is approximated by unity and incorporated into the polynomial $(1 - B)^{d_1+1}$.[3]

Let us now consider the stationary time series

$$b_t' = (1 - B)^{d_1+1}b_t \ . \qquad (3.16)$$

[3] In other words, we are substituting $\Phi(B)$ in (3.7) by $\Phi_2(B)\Phi_3(B)(1 - B)$.

The autocorrelation function $\{\rho_k\}$ for $\{b_t'\}$ is given by the solution to the equation

$$\Phi_1(B)\Phi_2(B)\Phi_3(B)\rho_k = 0 \ , \quad k > q_1 \ . \qquad (3.17)$$

If we let $\lambda_2 = be^{i\alpha}$ and $\lambda_3 = be^{-i\alpha}$ as in (2.7a) and (2.7b), then the solution to (3.17) is given by

$$\rho_k = A_{2,3}b^k \sin (k\alpha + \beta_{2,3}) + A_4\lambda_4{}^k + A_5\lambda_5{}^k + \ldots + A_{N+1}\lambda_{N+1}{}^k + C_1G_1{}^k + C_2G_2{}^k + \ldots + C_{p_1}G_{p_1}{}^k \ , \qquad k > q_1 - N - p_1 \ , \quad (3.18)$$

where $G_1{}^{-1}, G_2{}^{-1}, \ldots, G_{p_1}{}^{-1}$ are the roots of $\Phi_1(B) = 0$ and $A_{2,3}, \beta_{2,3}, A_4, A_5, \ldots, A_{N+1}$, and $C_1, C_2, \ldots, C_{p_1}$ are constants. (In practice, $N \gg q_1$ so that (3.18) is valid for $k \geq 0$.)

Since the eigenvalues λ_i, $i = 4, 5, \ldots, N + 1$, are smaller (in absolute value) than λ_2 and λ_3, their influence on the autocorrelation function $\{\rho_k\}$ will be negligible after a few lags (see (3.18)). Thus it would be desirable to eliminate these $N - 2$ eigenvalues from the model. It should be recalled that for a stationary and invertible ARIMA (p, d, q) model, the autocorrelation function follows a mixture of damped exponentials and/or damped sine waves after lag $(q - p)$, and that the first q values of the autocorrelation function depend on the parameters of the moving average polynomial (Box and Jenkins 1970, Section 3.4.2). Hence moving average parameters might be used to include the influence of the eigenvalues $\lambda_4, \lambda_5, \ldots, \lambda_{N+1}$ on the first few lags of the autocorrelation function, in which case these eigenvalues could be eliminated from the model. Thus a parsimonious version of the ARIMA model of (3.7) would be

$$\Phi_1(B)\Phi_2(B)(1 - B)^{d_1+1}b_t = \Theta_1(B)\Theta_2(B)a_t \ , \quad (3.19)$$

where $\Theta_2(B)$ is a moving average polynomial incorporated into the model to include the influence of the $N - 2$ eigenvalues $\lambda_4, \lambda_5, \ldots, \lambda_{N+1}$ on the autocorrelation function.

Since any stationary and invertible ARIMA model is fully determined by its autocorrelation function, and since the autocorrelation function of (3.19) is approximately equal to the autocorrelation function of (3.7), (3.19) represents an approximation to the ARIMA model of (3.7). As will be seen in the next section, we can obtain very good ARIMA models for the birth time series using up to five parameters.

3.3 Applications

Data for the female-birth time series for Norway for 1919–1974 were taken from various issues of the Statistical Yearbook of Norway. Various ARIMA models were tried, and the best encountered were an ARIMA (4, 1, 1) and an ARIMA (3, 1, 2). (See Table 1.) The criteria for choosing the best model(s) were:

(i) ability to incorporate the three largest eigenvalues λ_1, λ_2, and λ_3;
(ii) parsimony;
(iii) portmanteau lack-of-fit test (see Box and Jenkins (1970), ch. 8);

1. Estimated ARIMA *Models for Female-Birth Time Series for Norway, 1919–1974, in Thousands*

Model	*Fitted model*[a]	*Residual standard error*[b]	*Q*[c]	*Degrees of freedom*
ARIMA (4,1,1)	$(1 - .91B - .28B^2 + .16B^3 + .16B^4)\nabla b_t = (1 - .89B)a_t$ $(.37)\ (.39)\ (.30)\ (.20)\ \ (.24)$	1.126	16.06	25
ARIMA (3,1,2)	$(1 - 1.40B + .27B^2 + .21B^3)\nabla b_t = (1 - 1.36B + .44B)a_t$ $(.32)\ (.31)\ (.12)\ \ (.37)\ (.40)$	1.115	14.92	25

[a] Values in parentheses underneath coefficients are 95 percent confidence intervals.
[b] Residual standard error is an estimator for σ_a.
[c] Used for portmanteau lack-of-fit test. In both cases Q values are not significant when compared with χ^2 (25) distributions (which means that the models are adequate).

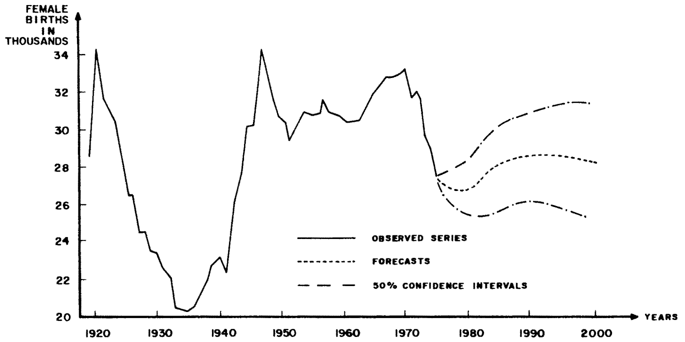
(iv) quality of forecasts, i.e., closeness of forecasts with available future values and size of confidence intervals.

Although various models with less than five parameters gave excellent results with respect to (ii), (iii), and (iv), only models with five parameters were able to incorporate the eigenvalues λ_2 and λ_3, i.e., the information on the length of generation.

The pair of largest complex roots for the autoregressive operators was calculated, and from them an estimator for the length of generation L for the period considered obtained. (See (2.8).) For the ARIMA (4, 1, 1) model, we found $L = 27$ years, and for the ARIMA (3, 1, 2) model, $L = 26$ years. A comparison between the values for the length of generation obtained from the ARIMA models with the length of generation calculated for various years during the period considered using classical methods is very favorable to the ARIMA models. According to Keyfitz and Flieger (1971), between the years 1950 and 1967 the length of generation varied from a maximum of 29.5 to a minimum of 27.2 years.

Forecasts for 1975 were calculated using both the ARIMA (4, 1, 1) and the ARIMA (3, 1, 2) models. Comparison with the number of female births for Norway during 1975 (27,496 births) shows that the forecasts for that year (28,567 for the ARIMA (4, 1, 1) and 28,728 for the ARIMA (3, 1, 2)) were very reasonable. Both forecasts overestimated the actual number of female births in 1975 due to the fertility decrease during that year.

Female-Birth Time Series for Norway, 1919–1975, Together with Forecasts and 50 Percent Confidence Intervals for Period 1976–2000[a]



[a] Forecasts and confidence intervals using ARIMA (4, 1, 1) model of Table 1. Source: Data from the *Statistical Yearbook of Norway* (1920–1975).

Forecasts for 1976–2000, together with 50 percent confidence intervals, were calculated (see the figure for the ARIMA (4, 1, 1) forecasts). The effect of the length of generation in the forecast function can be seen very clearly; it is responsible for the long damped sine wave which attains a minimum in 1978 and a maximum in 1992.

As mentioned in Section 3.1, it is also possible to obtain ARIMA models for *total* birth time series (female and male births considered together). To illustrate this point, the birth time series from Norway for the period 1919–1974 was considered, and the ARIMA (4, 1, 1) and ARIMA (3, 1, 2) models of Table 2 were obtained. It is interesting to note that the corresponding ARIMA models of Tables 1 and 2 are very similar. Their estimated autoregressive and moving average parameters are close to each other, and the residual standard errors for the ARIMA models of Table 2 are slightly bigger than two times the corresponding ones of Table 1 (2.334/1.126 = 2.07, 2.306/1.115 = 2.07). This should not constitute a surprise and is, indeed, an expected result. As mentioned at the end of Section 3.1, the ARIMA models for the total birth time series can be obtained by multiplying the ARIMA models of the female birth time series by unity plus the average sex ratio ≈ 2.05. Thus the ARIMA models for the total birth time series should have approximately the same autoregressive and moving average parameters as the ARIMA models for the female-birth time series, and its white noise terms $\{a_t\}$ should be approximately 2.05 times bigger than the white noise terms of the ARIMA models for the female-birth time series.

Forecasts for the Norwegian birth time series for 1976–2000, together with 50 percent confidence intervals, were obtained using both the ARIMA (4, 1, 1) and ARIMA (3, 1, 2) models (see Table 3 for the ARIMA (4, 1, 1) forecasts).

We also compared our forecasts with four alternative projections made by the Central Bureau of Statistics of Norway (Brunborg 1974). They vary from the highest alternative (I), in which fertility is kept constant, a little above the registered level in 1973, to the lowest alternative (IV), in which it is supposed that from 1977 fertility is approximately equal to the Swedish fertility of 1972. In all four alternatives, the sex ratio at birth is the one registered for 1962–1972, and the mortality rates are based on observations for 1968–1972 and are kept constant throughout the projection period. Our forecasts

2. *Estimated* ARIMA *Models for Total Birth Time Series from Norway 1919–1974 in Thousands*

Model	Fitted model[a]	Residual standard error[b]	Q[c]	Degrees of freedom
ARIMA (4,1,1)	$(1 - \underset{(.39)}{.88}B - \underset{(.40)}{.34}B^2 + \underset{(.30)}{.16}B^3 + \underset{(.20)}{.18}B^4)\nabla b_t = (1 - \underset{(.27)}{.91}B)a_t$	2.334	14.84	25
ARIMA (3,1,2)	$(1 - \underset{(.30)}{1.42}B + \underset{(.30)}{.29}B^2 + \underset{(.12)}{.22}B^3)\nabla b_t = (1 - \underset{(.35)}{1.39}B + \underset{(.38)}{.46}B^2)a_t$	2.306	15.19	25

[a] Values in parentheses underneath coefficients are 95 percent confidence intervals.
[b] Residual standard error is an estimator for σ_a.
[c] Used for portmanteau lack-of-fit test. In both cases Q values are not significant when compared with χ^2 (25) distributions.

for the ARIMA (4, 1, 1) model are approximately equal to their lowest alternative (IV). This seems to be a reasonable result considering the drop in fertility rates in Norway between 1973 and 1975.

3. *Forecasts for the Total Birth Time Series for Norway for 1976–2000 Using* ARIMA *(4,1,1) Model, in Thousands*

Year	50 percent lower limit	Forecast	50 percent upper limit
1976	53.891	55.479	57.067
77	52.198	54.416	56.634
78	50.977	54.007	57.037
79	50.250	54.021	57.792
1980	49.811	54.227	58.643
81	49.684	54.676	59.668
82	49.744	55.218	60.692
83	49.937	55.816	61.695
84	50.206	56.420	62.634
85	50.499	56.989	63.479
86	50.787	57.502	64.217
87	51.041	57.943	64.845
88	51.248	58.305	65.362
89	51.401	58.589	65.777
1990	51.495	58.798	66.101
91	51.537	58.940	66.343
92	51.529	59.025	66.521
93	51.481	59.062	66.643
94	51.398	59.062	66.726
95	51.291	59.035	66.779
96	51.165	58.989	66.813
97	51.028	58.932	66.836
98	50.886	58.870	66.854
99	50.743	58.809	66.875
2000	50.603	58.752	66.901

Two further observations are worth making. First, the difficulty of finding ARIMA models which incorporate λ_2 and λ_3 is a sign of the importance of short-term variations in birth time series. It should be mentioned that many of the estimated models with less than five parameters had autoregressive operators whose roots were complex conjugates. In most cases, such roots were responsible for fluctuations of periods of four or five years. A property of ARIMA models with autoregressive operators of order four or more is that they are able to incorporate both short-term and long-term cycles. The autoregressive operator for the ARIMA (4, 1, 1) model for Norway, e.g., has two pairs of complex roots, one responsible for the length of generation, and the other responsible for short-term fluctuations.

Second, a nice property of the ARIMA models is that their forecasts adjust very fast to sudden changes in the time series. If, e.g., there is a big drop in the number of births due to changes in the fertility pattern of the population, the ARIMA forecasts have a tendency to capture such changes quickly. This was verified when using data for the United States birth time series, and seeing how well estimated ARIMA models adjusted their forecasts to the decrease of births in the early seventies.

4. CONCLUSIONS

We saw how ARIMA models for birth time series can be obtained. The starting point was the discrete renewal equation for births in which the vital rates are assumed to vary with time, and migration is permitted. A parsimonious model was developed which incorporates the most important pieces of information including the length of generation and short-term variations in the vital rates. Forecasts were developed using Box and Jenkins time series techniques, and we saw how the length of generation influences such forecasts.

Perhaps the greatest advantage of our model is that the only data necessary are the time series of births for a medium period of time; fifty years seems to be enough. More values are desirable to obtain better estimates of the parameters but, at the same time, are undesirable since the longer the period considered the greater are the changes in the vital rates, so that a compromise is needed. Although we should not expect that every country would have such data available, in order to apply classical methods we usually need the vital rates for the various age-intervals as well as an assumption on how they will behave in the future.

The method was applied to Norway. Two models were obtained for which the length of generation compares very well with estimates of the length of generation for various years of the period considered. The difficulty in estimating a model which incorporates the information on the length of generation is another sign that one of the important sources of variation for birth time series is the short-term changes in the vital rates.

Our method can be naturally extended to include the population in the various age-intervals. Given a model for birth time series, we can express each age-interval group as a linear function of past births plus an error term, i.e., the population in age-interval j at time t can be expressed as the number of births during period of time $t - j - 1$ multiplied by the corresponding average survival rate plus an error term. Since total population is the sum of populations in the various age-intervals, total population can also be expressed as a linear function of past births plus an error term. The transfer function technique of Box and Jenkins might then be useful to forecast future populations for the various age-intervals taken separately or considered together.

[*Received July 1975. Revised November 1976.*]

REFERENCES

Bernadelli, H. (1941), "Population Waves," *Journal of Burma Research Society*, 31, 1–18.

Box, G.E.P., and Jenkins G.M. (1970), *Time Series Analysis Forecasting and Control*, San Francisco: Holden-Day.

Brunborg, H. (1974), "Population Projections for Norway 1973–2100 An Analytic Experiment," Oslo: Central Bureau of Statistics.

Goodman, L.A. (1967), "On the Reconciliation of Mathematical Theories of Population Growth," *Journal of the Royal Statistical Society*, Ser. A, 130, 541–53.

——— (1968), "Stochastic Models for the Population Growth of the Sexes," *Biometrika*, 55, 469–87.

Hiorns, R.W. (1972), "Mathematical Models in Demography," in *The Structure of Human Populations*, eds. G.A. Harrison and A.J. Boyce, Oxford: Clarendon Press, 110–27.

Harris, T.E. (1963), *The Theory of Branching Processes*, Berlin: Springer-Verlag.

Keyfitz, N. (1972), "Population Waves," in *Population Dynamics*, ed. T.N.E. Greville, New York: Academic Press, 1–38.

———, and Flieger, W. (1971), *Population-Facts and Methods of Demography*, San Francisco: W.H. Freeman & Co.

Le Bras, H. (1971), "Un Modèle de Projections Aleatoires," paper presented to the IUSSP Meeting in Mexico City, 1971.

Lee, R. (1974), "Forecasting Births in Post-Transition Populations," *Journal of the American Statistical Association*, 69, 607–17.

Leslie, P.H. (1945), "On the Use of Matrices in Population Mathematics," *Biometrika*, 33, 183–212.

———, (1948a), "Some Further Notes on the Use of Matrices in Population Mathematics," *Biometrika*, 35, 213–45.

———, (1948b), "On the Distribution in Time of the Births in Successive Generations," *Journal of the Royal Statistical Society*, Ser. A, 111, 44–53.

Lewis, E.G. (1942), "On the Generation and Growth of a Population," *Sankhyā*, 6, 93–6.

Norway, Central Bureau of Statistics (1920–1975), *Statistical Yearbook*.

Norway, Central Bureau of Statistics (1920–1975), *Vital Statistics and Migration Statistics*.

Pollard, J.H. (1966), "On the Use of Direct Matrix Product in Analyzing Certain Stochastic Population Models," *Biometrika*, 53, 397–415.

——— (1969), "A Discrete-Time Two-Sex Age-Specific Stochastic Population Incorporating Marriage," *Demography*, 6, 185–221.

——— (1970), "On Simple Approximate Calculations Appropriate to Populations with Random Growth Rates," *Theoretical Population Biology*, 1, 208–18.

Saboia, J.L.M. (1974), "Modeling and Forecasting Populations by Time Series— The Swedish Case," *Demography*, 11, 483–92.

Schweder, T. (1971), "The Precision of Population Projections Studied by Multiple Projection Methods," *Demography*, 8, 441–50.

Sykes, Z.M. (1969), "Some Stochastic Versions of the Matrix Model for Population Dynamics," *Journal of the American Statistical Association*, 64, 111–31.

Tukey, J.W. (1961), "Discussion Emphasizing the Connection Between Analysis of Variance and Spectrum Analysis," *Technometrics*, 3, 191–219.