

# A Statistical Look at Modeen's Forecast of the Population of Finland in 1934<sup>1</sup>

JUHA M. ALHO, Ph.D.

Professor

University of Joensuu, Department of Statistics

## Abstract

*Gunnar Modeen made the first cohort-component forecast for Finland in 1934. This was a time when demographic transition was just over, but that fact could not have been known at the time. Would it have made any difference if Modeen had had the tools of modern time-series analysis available? We find that the essential question of how to deal with changing trends would have still been difficult. However, the modern tools would have given the forecast user a realistic indication of the uncertainty of the forecast being made. This suggests that in developing countries that are undergoing transition now, more effort should be paid to the analysis of uncertainty of forecasting.*

**Keywords:** Box-Jenkins, forecast error, historical demography, judgmental methods, time series analysis

## Background of Modeen's forecast

Cannan's (1895) forecast for England and Wales was apparently the first cohort-component forecast. Several other cohort-component forecasts were made in the 1920s in Europe and the United States (DeGans 1999). In 1932, Gunnar Modeen, then actuary at the Central Statistical Office, had prepared a population forecast of the city of Helsinki. In 1934, the population statistics of the year 1930 and mortality statistics of the 1920s had been compiled and completed (Luther 1993). Consequently, Modeen prepared the first national cohort-component forecast for Finland, with 1930 as the initial year (Modeen 1934a). The forecast extended to the year 1980, and predicted that the population would peak at just under four million during the 1970s. However, the actual population was 4.8 million at that time. Thus, the error incurred during a 50 year forecast period was about 17 percent.

<sup>1</sup>An early version of the paper was presented at the Workshop on Historical Demography, in Rostock, in May 1999.

The question that we will address is whether it would have made any difference in the forecast if modern statistical time-series methods had been available to Modeen. The issue is of considerable methodological interest since, at the turn of the millennium, many developing countries are facing the demographic transition. The success of future forecasts of world populations largely depends on how the timing and magnitude of the transition is handled. Our main finding is that there has been little progress in the ability to forecast changes in trends. Nonetheless, time-series techniques are able to provide a realistic assessment of the uncertainty of forecasting.

As a technical exercise, Modeen's efforts were comparable to similar research done elsewhere. He acknowledges the works of Bowley (1924), Greenwood (1925), Gini (1930), Wicksell (1926, 1934), Jensen (1930), and Sauvy (1932) as providing a suitable methodology for the task. Modeen criticized, however, the logistic model introduced by Verhulst (1838), and later popularized by Pearl and Reed (1920) and Yule (1925), as being untenable (Modeen 1934b). He argued that the main defect of the logistic and exponential curves was that they predicted permanent growth (or decline). No change from growth to decline or vice versa was possible under these techniques. The cohort-component methodology, on the other hand, had no such limitation.

Finland has continuous data series for several demographic characteristics: 1) population size by age and sex, starting in 1749; 2) age-specific mortality, starting in 1751; and 3) age-specific fertility, starting in 1776 (Nieminen 1999). This is plenty of data necessary for the usual time-series techniques. However, we may not be able to perfectly replicate Modeen's work for several reasons. First, he did a complete cohort-component forecast, but redoing it would be too laborious. We will only handle total population via growth rates. Second, we will use data compiled by Turpeinen (1978). As a result of later revisions the data series are better than the data available to Modeen, although the main features of the demographic development were known to him. The series end at 1925, so it does not exactly match the jump-off time of Modeen's forecast. Third, since we know the demographic trends since 1934, this may have influenced our modeling. The reader must judge to what extent this has occurred.

In Section 2, we define empirical measures that capture the major developments in the growth rates. In Section 3, we present empirical data on past birth and death rates. Section 4 presents statistical forecasts of future vital rates and assesses the uncertainty they imply for the total population. In the final section, we suggest that the use of modern statistical methods might have changed the forecasts in the 1930s.

## Quantifying forecastability

Our goal is to characterize the unpredictable variation of fertility and mortality in Finland before the jump-off year of Modeen's forecast, the year 1930. We will do this by decomposing crude fertility and mortality multiplicatively into a level effect and an effect due to the age composition. This will enable the analysis of uncertainty to be executed using level effects as summary measures, without having to conduct full cohort-component calculations.

Let  $V(x,t)$  be the number of people in age-sex group  $x$ , in the beginning of year  $t$ . Define the corresponding total population  $V(t) = \sum_x V(x,t)$ , and age distribution  $v(x,t) = V(x,t)/V(t)$ . Define  $\eta_1(t) = \sum_x h_1(x,t)v(x,t)$ ,  $i = 1, 2$ . Letting  $h_1(x,t)$  be the expected number of births per person in age-sex group  $x$  ( $h_1(x,t) = 0$ , where  $x$  corresponds to males and to females outside childbearing ages), and  $\eta_1(t)$  is the crude birth rate during time  $t$ . Defining  $h_2(x,t)$  as the death rate of age-sex group  $x$ ,  $\eta_2(t)$  is the crude death rate of the year  $t$ . Assuming the rates to be constant during year  $t$  and migration to be zero, we obtain the recursion

$$V(t+1) = \exp(\eta_1(t) - \eta_2(t))V(t),$$

where  $\eta_1(t) - \eta_2(t)$  is the growth rate.

Suppose we have two standard age distributions  $v_i(x)$ ,  $i = 1, 2$ , and let  $\xi_i(t)$  be the corresponding (directly) standardized birth rate for  $i = 1$  and survival rate for  $i = 2$ . We then have  $\eta_1(x,t) = \xi_1(t)c_1(t)$ , where

$$c_i(t) = \sum_x h_i(x,t)v(x,t)/\sum_x h_i(x,t)v_i(x),$$

"corrects" for the effect of the standardization. Therefore, the growth rate is represented as

$$\eta_1(t) - \eta_2(t) = \xi_1(t)c_1(t) - \xi_2(t)c_2(t).$$

Letting  $(\xi_i, c_i)$  be the point of averages, the following are the representations

$$\xi_i(t)c_i(t) = \xi_i c_i + c_i(\xi_i(t) - \xi_i) + \xi_i(c_i(t) - c_i) + (\xi_i(t) - \xi_i)(c_i(t) - c_i)$$

consisting of an overall mean  $\xi_i c_i$ ; two main effects  $c_i(\xi_i(t) - \xi_i)$  and  $\xi_i(c_i(t) - c_i)$ ; and an interaction term  $(\xi_i(t) - \xi_i)(c_i(t) - c_i)$ . Applying these to the growth rate, the result is

$$\eta_1(t) - \eta_2(t) = \rho + c_1(\xi_1(t) - \xi_1) - c_2(\xi_2(t) - \xi_2) + H(t) + R(t),$$

where  $\rho = \xi_1 c_1 - \xi_2 c_2$  is a measure of the level of past growth,  $H(t) = \xi_1(c_1(t) - c_1) - \xi_2(c_2(t) - c_2)$  is a factor determined by the age structure of fertility and mortality, and  $R(t) = (\xi_1(t) - \xi_1)(c_1(t) - c_1) - (\xi_2(t) - \xi_2)(c_2(t) - c_2)$  is an interaction term.

Full cohort-component calculations provide information nearly equivalent to  $H(t)$ , since  $H(t)$  depends on the future levels of fertility and mortality. This occurs via the future age-distribution of the population and of the rates, but not their absolute levels. The term  $R(t)$  is of a smaller order of magnitude than the other terms, because the coefficients of variation of the  $c_i(t)$  are small. It follows that the accuracy of a cohort component forecast is determined by the unpredictable variation of the standardized measures  $\xi_i(t)$ . In the following analysis we will concentrate on the standardized measures and their logarithms  $\rho_i(t) = \log(\xi_i(t))$ .

## Prior evidence on level and composition

Based on data published by Turpeinen (1978), we will analyze Finland's mortality and fertility rates until the year 1925. They consist of estimates of (female) age-specific fertility by five year age groups (15-19, ..., 45-49) since 1776, and estimates of age-specific mortality (both sexes combined) by five year age groups (0-4, ..., 60-64, 65+) since 1751.

The standard age schedules were chosen so that  $v_1(x)$  is a constant for female ages 15-49 and zero elsewhere. In other words, our *standardized birth rate* (SBR) is the total fertility rate divided by 35. To calculate the *standardized death rate* (SDR), the schedule  $v_2(x)$  is equivalent to the age distribution of the male life table population of the years 1921 to 1930 (Kannisto and Nieminen 1996).

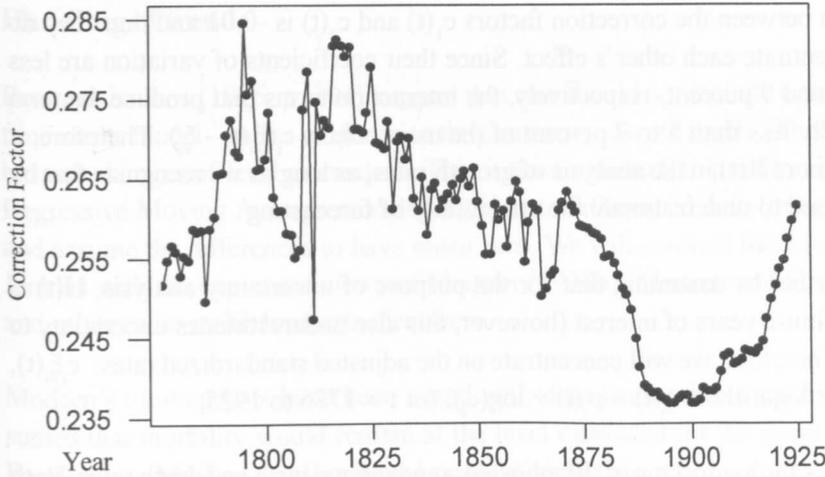
In the following section, we will analyze birth rates and death rates. We then summarize the results by examining the growth rates in Section 3.2.

### Fertility and mortality until 1925

There was a 61 percent overall decline of the *crude birth rate* (CBR), from 0.039 in 1776 to 0.024 in 1925. There was an almost equivalent decline in the SBR, of 58 percent. Figure 1 provides the correction factors  $c_1(t)$  for fertility. They have been calculated from the ratio CBR/SBR. Although the correction factors form a seemingly dramatic curve, the coefficient of variation of  $c_1(t)$  is only .046. In other words, the average variability of  $c_1(t)$  is less than 5 percent from the mean level.

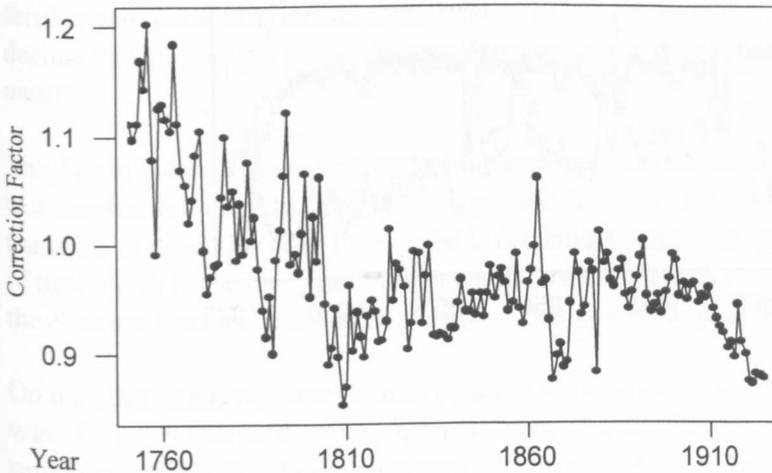
A formal estimate of the level effect given by the SBR and correction factors  $c_1(t)$  on the CBR is obtained by regressing  $\log(\eta_1(t))$  on  $\rho_1(t)$ . We find that  $R^2 = 0.88$ , which means that the SBR explains 88 percent of the relative variability of the CBR, from 1776 to 1925.

**Figure 1.** Correction factors for fertility



The decline of *crude death rate* (CDR) from about 0.030 in 1751 to 0.015 in 1925 corresponds to a 50 percent decline. The SBR decreases less, approximately 37 percent, from about 0.027 to about 0.017. From Figure 2, we find that the correction factors alone would have decreased CDR by about 20 percent. Through the multiplication  $0.63 \times 0.80 = 0.50$  we obtain the observed decline. When we regress  $\log(\eta_2(t))$  on  $\rho_2(t)$ , then  $R^2 = 0.93$ , which is higher than the explained variation for fertility. Finally, the coefficient of variation of  $c_2(t)$  is .068. In other words, the average variation is less than 7 percent from the mean.

**Figure 2.** Correction factors for mortality



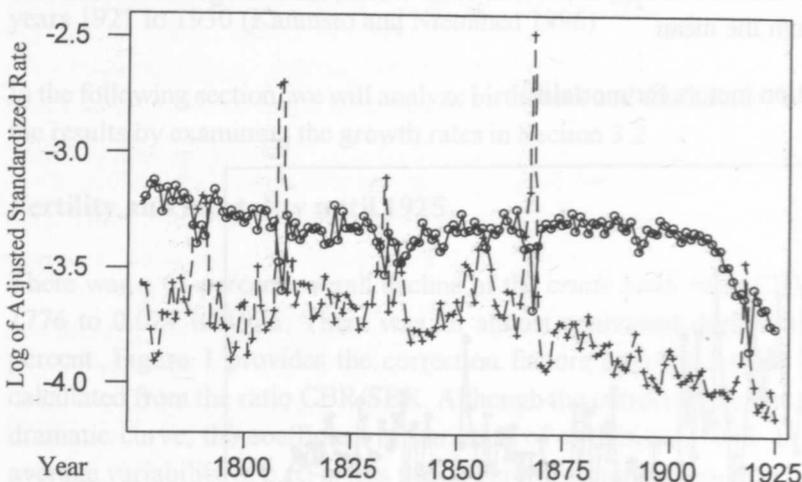
## Growth rates in 1776-1925

The correlation between the correction factors  $c_1(t)$  and  $c_2(t)$  is  $-0.01$  and thus they do not tend to accentuate each other's effect. Since their coefficients of variation are less than 5 percent and 7 percent, respectively, the interaction terms that produce the term  $R(t)$  are typically less than 5 to 7 percent of the main effects  $c_i(\xi_i(t) - \xi_i)$ . Therefore, it is justified to ignore  $R(t)$  in the analysis of growth rates, as long as we recognize that by doing this we tend to underestimate the uncertainty of forecasting.

We simplify further by assuming that for the purpose of uncertainty analysis,  $H(t)$  is known for the future years of interest (however, this also underestimates uncertainty to some extent). Therefore, we will concentrate on the adjusted standardized rates:  $c_i \xi_i(t)$ ,  $i = 1, 2$ , or their logarithms  $\zeta_i(t) = \rho_i(t) + \log(c_i)$ , for  $t = 1776$  to 1925.

Figure 3 displays the logarithms of the adjusted standardized birth and death rates. Both series show a declining trend. During the entire period from 1776 to 1925, a linear regression on time maintains the slope  $-0.0015$  for the birth rate and slope  $-0.0028$  for the death rate. However, towards the end of the period both declines accelerate. A key issue in the statistical forecasting of the level of fertility and mortality, therefore, concerns how these trends are handled or figured into the model.

**Figure 3.** Logarithm of the adjusted standardized birth rates (o) and death rates (+).



## Modeling fertility, mortality, and population growth

### Handling of trends

Box and Jenkins (1976, p. 194) suggest that one should not include a trend term into a time series model "unless evidence to the contrary presents itself." Their alternative is to treat changes of level as random. This is accomplished, in practice, by fitting Auto Regressive Moving Average models (ARMA) to differenced data (one or more times), and assume the differences to have mean zero. We will contrast the advice of Box and Jenkins with the expressed views of Modeen (1934b), Modeen and Fougstedt (1938), and others who make forecasts elsewhere.

Modeen's assumptions have been tested with data from other countries. Modeen assumed that mortality would remain at the level estimated for the years 1921 to 1930. For fertility, Modeen's assumption I postulated that the absolute number of births remains constant. Wicksell (1934), in a forecast for Norway, as well as the German Statistisches Reichsamtsamt (1930), used such assumptions also. Modeen's assumption II postulated that the general fertility rate would remain at the average level of the years 1931 and 1932. However, both Wicksell and the Reichsamtsamt considered changes of level. Modeen assumed net migration to be zero, as did Wicksell and the Reichsamtsamt.

Although Modeen considered his calculations as being somewhat hypothetical, the next Finnish national forecast prepared by Modeen and Fougstedt (1938) also assumed that mortality would remain at a constant level (1931 to 1935). Even if a further decline were likely, they argued that "it is very difficult to predict, how big this decrease is in different age groups". The three assumptions concerning births were also similar to the earlier ones: 1) either the total number of births was assumed to remain constant; or 2) fertility was assumed to remain at the 1931 to 1935 level; or 3) fertility was assumed to decline by 1 percent per year, within each age category. Again, zero net migration was assumed.

The data in Figure 3 shows that, apart from peaks of short duration, mortality remained at a constant level during the first century of the data period. Subsequently, apart from the effect of the civil war in 1918, a stable declining trend began around 1870. Neither of the Finnish forecasts considered the possibility that the trend might continue despite the evidence from the past half of a century. This is a bit surprising.

On the other hand, we observe from Figure 3 the final accelerated decline for fertility was of a much shorter duration, approximately 25 years. When writing about the assumptions of the first Finnish forecast, Modeen (1934b, p. 361-362) referred to a French forecast in which one variant had assumed fertility rates to decline to the level

of the region of Seine, "a fairly urbanized area". However, Modeen assumed that this and other such hypotheses were based more on "speculation" rather than his own assumptions of constant fertility.

Modeen and Fougstedt (1938, p. 8) also argued that the urbanization of society will most likely lead to a decrease in both marriage rates and in marital fertility. This was the prevailing view elsewhere. In Germany, Burgdörfer (1932, p. 32) wrote of Berlin as the "unfruchtbare Stadt" and worried about the unhappy consequences of the "Zweikindersystem". In Sweden, Myrdal and Myrdal (1934, pp. 87-88, 94) attributed the decline in fertility rates to improved contraception as well as secularized rationality that follow from urbanization. They suggested that the decline would continue for the "nearest decades". Whelpton (1947, pp. 28-29) argued similarly that U.S. fertility rates would continue to decline as a result of urbanization, women's increased labor force participation, and improved contraception.

Clearly, Modeen had a basis for assuming that the trends would continue to decline, but he was uncertain of the speed and duration of the decline. A statistical reconciliation of the two points of view may be given by considering the fact that modeling error, or bias, is a major source of uncertainty in forecasting (Alho and Spencer 1997). In this case, bias could mean, for example, that the more or less linear decline in fertility would only seem to be linear because the possible curvature of the trend had not manifested itself yet. In some cases, it is possible to set empirical bounds on the bias, but this may require complex analyses (Alho and Spencer 1985). On the other hand, the advice of Box and Jenkins may be interpreted as providing an automated way to treat bias in probabilistic terms: if one sees evidence of a trend but deliberately ignores it, the missing trend component will be part of the residual error; and, as such, it will be incorporated into the prediction intervals of the future values. However, a potential drawback is that this may sometimes result in overly conservative prediction intervals (i.e. too wide)(Alho 1998, p. 17).

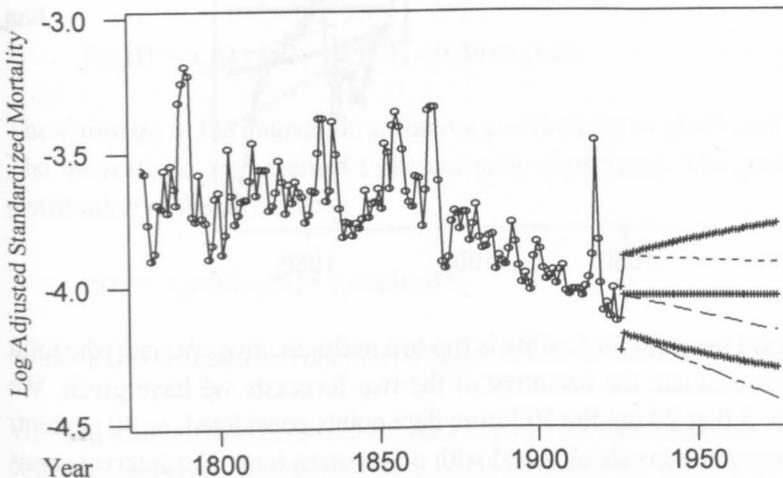
### **Statistical aspects of fertility and mortality**

Let us consider mortality first. Not wanting the three largest peaks to dominate estimates of residual error, we replaced them with previous values. In all, five annual values were changed. The autocorrelation function of the series moves to zero very slowly, confirming the visual impression of nonstationarity. The autocorrelation function of the differenced series is nearly flat, with  $-0.28$  at lag = 2 as the biggest deviation from zero. The partial autocorrelation has also a peak at lag = 2. This suggests that there may be a moving average component of order 2 in the series. Indeed, experiments with ARIMA( $p, 1, q$ ) models (i.e., ARMA( $p, q$ ) models applied to the first differences of the data series) with  $p + q \leq 2$  show that ARIMA(0,1,2) fits the best. A Box-Pierce goodness-of-fit statistic demonstrates that the model fits well. While adding a constant

term to the model improves the fit slightly, the estimated constant is not significantly different from zero. From a statistical point of view, one is well justified in following the advice of Box and Jenkins.

Figure 4 presents the original data series (with the peaks cut) and two forecasts with 67 percent prediction intervals. The intervals marked with dashed lines are based on an  $ARIMA(0, 1, 2)$  model without a constant term. After an initial jump, the forecast function is a constant. The dashed lines represent a forecast with a constant term. The prediction interval for the model with a constant reflects the assumption that the downward trend is real.

**Figure 4.** 67 percent interval forecast of mortality with a constant (-) and without a constant (+)

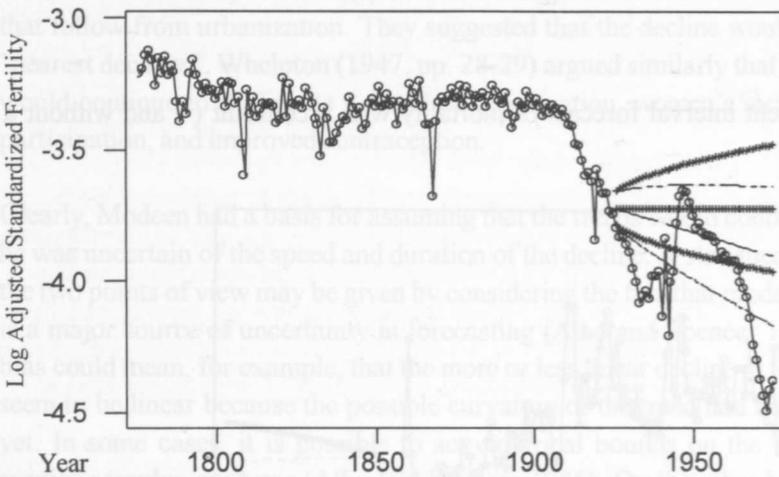


Let us consider fertility now. The autocorrelation function of the fertility series of the years 1776 to 1925 does not conclusively suggest nonstationarity. An  $AR(1)$  process with the first autocorrelation .75 could produce a similar autocorrelation function. This would be supported by the partial autocorrelation function, as well. However, it is clear from the graph that the end of the series does not result from a stationary process. Consider the first differences then. Their autocorrelation function has a negative peak of  $-.38$  at lag = 1. The other values are much smaller. The second partial autocorrelation is  $-.20$ . Either an  $AR(1)$  model or a  $MA(2)$  model might provide a reasonable fit. Experimenting with  $ARIMA(p, 1, q)$  with  $p + q \leq 2$  showed that  $ARIMA(0, 1, 1)$  provided as good a fit as the more complex models. It also was well fitted according to the Box-Pierce statistics. This model also provides the basis of the so called exponentially weighted moving average procedure that might be used as a model in its own right (Chatfield 1996, p. 70). As was the case for mortality, the constant term is not signifi-

cantly different from zero. Hence, the advice of Box and Jenkins finds support in the data analysis.

Figure 5 shows how the inclusion of the constant influences both the point forecast and the 67 percent prediction intervals. The situation is almost the same as for mortality. Model choice has a strong impact.

**Figure 5.** 67 percent interval forecasts of fertility and the actual development



Since our standardized measure for fertility is (up to a multiplicative constant) the total fertility rate, we can evaluate the accuracy of the two forecasts we have given. We observe from Figure 5 that 20 out of the 50 future data points considered, or 40 percent, are within the 67 percent intervals obtained with no constant term. For intervals constructed with a constant term, the actual coverage probability is 48 percent. These are not exceptionally poor values since the future data are highly autocorrelated, i.e. most points tend to be on one side of the forecast.

Recall also that 95 percent intervals are almost exactly twice as wide as the 67 percent intervals. In this case, a forecast user would have done well to prepare for contingencies implied by the 95 percent limits. Further, in the case of fertility forecasting with jump-off year 1925, the forecaster would initially have done better in following the theories of urbanization rather than the pragmatic advice of Box and Jenkins. However, the situation changed after the early 1930s, when the fertility decline halted. In fact, as we have noted above, Whelpton assumed as late as 1947 that the U.S. fertility rate would continue to decline. At that time, the advice of Box and Jenkins would have produced better results.

## Uncertainty in the forecast of the total population

We analyze the uncertainty of the future growth rates based on estimates obtained from adjusted standardized birth rates  $c_1 \exp(\rho_1(t))$  and adjusted standardized death rates  $c_2 \exp(\rho_2(t))$ , where  $c_1 = .26$  and  $c_2 = .97$ . We ignore the contributions of  $H(t)$  and  $R(t)$  (as defined in Section 2) to the uncertainty. The cross correlation of the first differences of the logarithms of the standardized rates is  $-.43$ . Wars and crop failures of the past seem to explain the high negative value. In an industrial society, such effects are expected to be smaller. Thus, for the purpose of analyzing the accuracy of the forecast in 1930 to 1980 we will assume the correlation to be zero.

Based on the ARIMA analyses of Section 4.2 we assume the following moving average models for the first differences

$$\nabla \rho_1(t) = a_1(t) - 0.51 \times a_1(t-1)$$

and

$$\nabla \rho_2(t) = a_2(t) - 0.23 \times a_2(t-1) - 0.49 \times a_2(t-2).$$

The estimates of the innovation variances are  $\text{Var}(a_1(t)) = .0049$  and  $\text{Var}(a_2(t)) = .0142$ . The innovations are assumed to be normally distributed. The goal is to estimate the distribution of the difference

$$\eta(t) = c_1 \exp(\rho_1(t)) - c_2 \exp(\rho_2(t)).$$

Since  $\eta(t)$  is a nonlinear function of  $\rho_i(t)$ , we use simulation for estimation.

We used Minitab to generate 3,000 sample paths for the difference. Accepting Modeen's cohort-component calculations as a point forecast, the relative error he should have expected for a  $t$ -year forecast, would have been the same as the uncertainty in  $\exp(\eta(0) + \dots + \eta(t-1))$ . Table 1 compares the forecast, and the corresponding 67 percent forecast interval to the observed development.

**Table 1.** Modeen's Forecast of the Total Population of Finland, Upper and Lower Endpoints of 67 percent Prediction Intervals, and the Actual Population Size (in Thousands), for the Years 1940 to 1980.

Year	Forecast	Lower 67 %	Upper 67 %	Actual
1940	3657.9	3569.8	3750.5	3695.6
1950	3840.4	3627.8	4090.3	4029.8
1960	3950.4	3554.0	4384.8	4446.2
1970	3993.5	3433.9	4694.7	4598.3
1980	3986.8	3233.0	4969.6	4787.8

We find that apart from the time around year 1960, the actual population size was within the 67 percent prediction intervals. Note, moreover, that the 95 percent intervals would have been approximately twice as wide. In this case, at least, even the probabilistic 67 percent prediction intervals would have given a realistic indication of the uncertainty to be expected.

## Discussion

We have discussed the historical background of Modeen's forecast. In Finland, and elsewhere, it was generally thought that due to the urbanization of society, fertility would continue to decline and remain low. Mortality was also thought to decline, but Modeen was reluctant to speculate about the speed of this decline.

Statistical analyses using standard ARIMA techniques lead to the same dilemma. Since the series being analyzed are nonstationary, the key decision one has to make in forecasting is whether or not to include a constant term into the model. Although the advice of Box and Jenkins, both among the most eminent statisticians of the century, is that the constant term probably should not be used, this is a decision that cannot be made on statistical grounds alone.

In Modeen's case, the advice is supported by the fact that the estimate of the constant term is not statistically significantly different from zero. A forecaster reluctant to make either decision can always act conservatively. In this case, he or she could take the possible modeling error into account by using the maximum of the upper prediction intervals, and the minimum of the lower prediction intervals, in inference. A less drastic approach would give an equal weight to both models and interpret the resulting prediction intervals in Bayesian terms (cf., Draper 1995). The best forecast would then be the mean of the two point forecasts, and the 67 percent intervals would be somewhat closer to the mean than the present two outermost limits.

A new development in the methodology of population forecasting has been the probabilistic handling of forecast uncertainty. A simple analysis based on the standardized birth and death rates can be used to develop approximate prediction intervals for the total population. Even though a number of simplifying assumptions are made that tend to underestimate uncertainty, we find that the estimates would have given Modeen a realistic indication about the uncertainty of forecasting from 1930 to 1980 in Finland. This approach is feasible for many developing countries today.

## References

- Alho, J.M. 1998. A stochastic forecast of the population of Finland. Reviews 1998/4. Helsinki: Statistics Finland.
- Alho, J.M. and B.D. Spencer. 1985. Uncertain population forecasting. *Journal of the American Statistical Association*, 80, 306-314.
- Alho, J.M. and B.D. Spencer. 1997. The practical specification of the expected error of population forecasts. *Journal of Official Statistics*, 13, 203-226.
- Box, G.E.P. and G.M. Jenkins 1976. Time series analysis, 2<sup>nd</sup> ed. San Francisco: Holden-Day.
- Bowley, A.L. 1924. Births and population of Great Britain. *Journal of the Royal Economic Society*, 34, 188-192.
- Burgdörfer, F. 1932. Volk ohne Jugend. Berlin: Kurt Vowinkel.
- Cannan, E. 1895. The probability of cessation of the growth of the population in England and Wales during the next century. *The Economic Journal*, 5, 505-515.
- Chatfield, C. 1996. The analysis of time series, 5<sup>th</sup> ed. London: Chapman & Hall.
- DeGans, H.A. 1999. Population forecasting 1895-1945. Dordrecht: Kluwer.
- Draper, D. 1995. Assessment and propagation of model uncertainty (with discussion). *Journal of the Royal Statistical Society B*, 57, 45-97.
- Gini, C. 1930. Calcolo di previsione della popolazione italiana del 1921 al 1961. *Notiziario demografico* 1930, 8-9.
- Jensen, A. 1930. Den danske befolkningens horoskop. *Nationaløkonomisk tidsskrift* 1930, 13-226.
- Greenwood, M. 1925. The growth of population in England and Wales. *Metron*, 5, 66-85.
- Kannisto, V. and M. Nieminen. 1992. *Revised life tables for Finland 1881-1990*. Population 1996:2. Helsinki: Statistics Finland.
- Luther, G. 1993. Suomen tilastotoimen historia vuoteen 1970. Helsinki: Tilastokeskus.
- Modeen, G. 1934a. Suomen väkiluvun vastainen kehitys. Tilastokatsauksia n:o 10. Helsinki: Tilastokeskus.
- Modeen, G. 1934b. Suomen väkiluvun tuleva kehitys ja sen taloudelliset seuraukset. *Kansantaloudellinen aikakauskirja*, VI, 351-378.
- Modeen, G. and G. Fougstedt 1938. Laskelmia Suomen vastaisesta väestönkehityksestä. Sarja VI, Väestötilastoa 89. Helsinki: Tilastokeskus.
- Myrdal, A. and G. Myrdal 1934. Kris i befolkningsfrågan. Stockholm: Bonniers.
- Nieminen, M. 1999. *Väestötilastoja 250 vuotta*. Väestö 1999:8. Helsinki: Tilastokeskus.
- Pearl, R. and L.J. Reed. 1920. On the rate of growth of the population of the United States since 1790 and its mathematical representation. *Proceedings of the National Academy of Science* 6, 275-288.
- Thompson, W.S. and P.K. Whelpton. 1933. Population trends in the United States. New York: McGraw-Hill.
- Sauvy, A. 1932. Calculs démographiques sur la population française jusqu'en 1980. *Journal de la société de statistique de Paris*, 1932, 338-347.
- Statistische Reichsamt 1930. Volkszählung. Statistik des Deutschen Reichs, Band 401,II., Berlin.
- Turpeinen, O. 1978. Fertility and mortality in Finland since 1751. *Population Studies*, 33, 101-114.
- Verhulst, P.F. 1838. Notice sur la loi que la population suit dans son accroissement.

*Correspondence Mathématique et Physique Publiée par A. Quételet (Bruxelles) 10, 113-121.*

Whelpton, P.K., H.T. Eldridge and J.S. Siegel. 1947. Forecasts of the population of the United States 1945-1975. Washington D.C.: U.S. Bureau of the Census.

Wicksell, S.D. 1926. Sveriges framtida befolkning under olika förutsättningar. *Ekonomisk Tidskrift*, 28, 91-123.

Wicksell, S.D. 1934. Bidrag till den formella befolkningsteorien. Med tillämpningar på Norges befolkning. *Statsökonomisk Tidskrift* 1934, Hefte 1-2, 1-94.

Yule, G.U. 1925. The growth of population and the factors which control it. *Journal of the Royal Statistical Society* 86, 1-58.