The New Zealand Statistician 21(1)

May 1986

pp 35-40

POPULATION PROJECTIONS FOR AUSTRALIA AND NEW ZEALAND BY THE LOGISTIC METHOD

James A. Koziol
Department of Basic and Clinical Research,
Scripps Clinic and Research Foundation,
La Jolla, California 92037,
United States of America
and
Henry C. Tuckwell*
Department of Mathematics,
Monash University,
Clayton, Victoria 3168,
Australia

The use of the logistic curve as a model for the populations of Australia and New Zealand is investigated. It is found that Australia's population data may be satisfactorily fitted with two logistic curves, one for data prior to 1947, and the other for subsequent data. On the basis of the latter curve, the asymptotic limit predicted for Australia's population is 24.5 million. A similar bifurcation is found with the logistic fit of New Zealand's population data and a 1986 population of 3.374 million is predicted for New Zealand.

INTRODUCTION

There are two widely used deterministic methods for population projection (see, for example, Keyfitz, 1968). These are the *component method* and the use of the *logistic growth curve.*

The logistic model is phenomenological in that the logistic differential equation was not derived from first principles, but instead was postulated as one of the simplest first order differential equations whose solutions would approach a finite limit at infinite times (Verhulst, 1983). Nevertheless, it was found to fit rather precisely the population data of many European countries. Recently, Leach (1981) showed that the logistic curve model could be satisfactorily applied to population data for Scotland, England, and the United States. He further argued that a logistic or similar curve can be used to predict population growth, because it can subsume the disparate elements of birth, death, and migration in arriving at an accurate representation of population trends.

In this note we apply the logistic model to population data for Australia and New Zealand, and find that the data from each country may be fitted satisfactorily with two logistic curves. Interestingly, the transitions from one curve to another occur in a time span which coincides with that found for a similar transition in the United States population data by Leach. This suggests that the population dynamics of Australia and New Zealand and of the United States have been similar.

* To whom correspondence should be addressed.

METHODS AND RESULTS

Solutions of the logistic differential equation may be written

$$N(t) = \frac{K}{1 + (K/N_0 - 1)\exp(-rt)},$$

where N(t) is the population size at time t, $N(0) = N_0$, K is the upper limit or carrying capacity, and r is the growth rate parameter.

Given population data for Australia at various time points, we used the criterion of least squares to find estimates of the three unknown parameters N_0, r and K which yielded the best fitting logistic curve. To this end we used an IMSL numerical routine that implemented a nonlinear least squares fit using a finite difference Levenberg-Marquardt algorithm (Levenberg, 1944; Marquardt, 1963). Computations were performed in double precision on the Monash University VAX 11/780 computer system.

The data employed were obtained from the Australian Year Book (Cameron, 1979), and are given in Table 1; parameter estimates are given in Table 2.

Table 1
Population data for Australia: 1881-1976

Year	Population (in millions)
1881	2.2502
1891	3.1778
1901	3.7738
1911	4.4550
1921	5.4357
1933	6.6298
1947	7.5794
1954	8.9865
1961	10.5483
1966	11.5995
1971	12.9372
1976	13.9155

Source: R. J. Cameron, Ed. (1979). *Year Book Australia.* Australian Bureau of Statistics, Canberra, Australia.

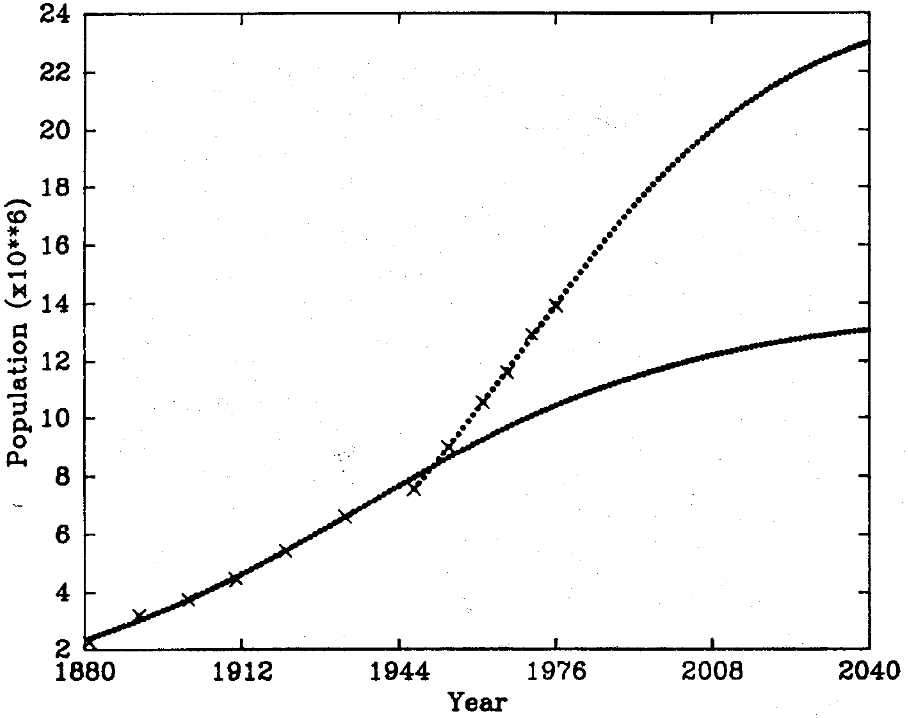
Table 2
Estimates of N_0, K and r for Australia

Census dates	**N(0)**	**K**	**r**	**Residual sum of squares**
1881-1933	2.389	13.733	.0285	6.11
1947-1976	7.565	24.527	.0375	2.14

N.B. N_0 and K are x10**6; r is x(1/yr); the residual sum of squares RSS is x10**10. Models fitted by an IMSL Levenberg-Marquardt algorithm routine.

We used population data at twelve different time points, as in Table 1. We initially fitted a logistic curve to the first six, and them emulated Leach's procedure of successively increasing the number of data points. We soon found, however, that the fits became progressively worse, as indicated by unacceptably large values for the norm of the gradient. We then discovered that two logistic curves would give a much more precise representation of the data; a changepoint between the two curves after 6 data points yielded the best fit, in terms of minimizing the total residual sums of squares about the curves. A graph of the population data from Table 1, as well as the two logistic curves indicated in Table 2, is given in Figure 1. There is a clear transition from one curve, fitting the data well for 1881-1933, to the other curve, from 1947 onward. Leach found a similar transition in the United States population data between the 1940 and 1950 census figures.

Figure 1. Australian Population Growth



The editor has kindly undertaken a similar analysis of New Zealand population data, which we now report. The data employed by him were obtained from the *New Zealand Official Yearbook* (1975, 1984), and are listed in Table 3; his parameter estimates (obtained using GENSTAT OPTIMIZE) are given in Table 4.

Table 3
Population data for New Zealand: 1901-1981

Year	Population (in millions)
1901	0.815862
1906	0.936309
1911	1.058312
1916	1.149225
1921	1.271668
1926	1.408139
1936	1.573812
1945	1.702330
1951	1.939472
1956	2.174062
1961	2.414984
1966	2.676919
1971	2.862631
1976	3.129383
1981	3.175737

N.B.Members of New Zealand Armed Forces overseas excluded.

Sources: Department of Statistics (1975). *New Zealand Official Yearbook.* Government Printer, Wellington, New Zealand.
Department of Statistics (1984). *New Zealand Official Yearbook.* Government Printer, Wellington New Zealand.

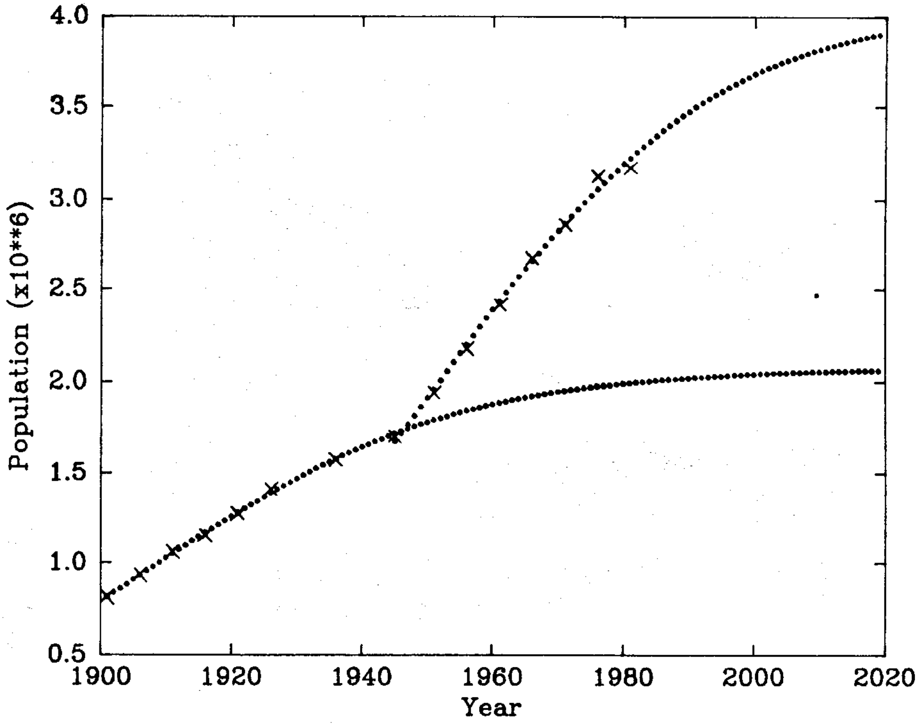
Table 4
Estimates of N_0, K and r for New Zealand

Census Dates	N_0	N(initial)	K	r	RSS
1901-1936	.795	.818	2.077	.0453	.0893
1945-1981	.312	1.673	4.070	.0473	1.02

N.B. N_0, N(initial) and K are x10**6; r is x(1/yr); RSS is x10*10. Time zero is 1900, initial time is 1901 for the pre-war set and 1945 for the post-war set. Models fitted by GENSTAT OPTIMIZE.

As is shown in Figure 2, the pre- and post-war data fall naturally along two logistic curves, although the exact point of transition is somewhat ill-defined.

Figure 2. New Zealand Population Growth



The editor has thus raised the important issue of whether the pre- and post-war census figures should be fitted separately, as compared with fitting them jointly and simultaneously estimating the changeover or join point. In the linear regression context, the analogous problem of switching regressions has received widespread attention, and most suggested procedures proceed to estimate all parameters jointly. Two principal approaches to the switching regressions problem may be termed parametric [*e.g.*, Quandt's (1958, 1960) log-likelihood ratio technique, or Hinkley (1969)], and Bayesian [e.g. Bacon and Watts (1971), or Choy and Broemeling (1980)]. Generalizations of these approaches to non-linear models is not so straightforward, however, particularly if one's goal is appropriate distribution theory for estimation and inference. In this regard, Leach's technique of estimating the join point at which the switch from one logistic equation to another has occurred [by minimizing (over the various cut-points) the total residual sums of squares about the two lines] may be viewed as a least-squares approximation to Quandt's maximum likelihood technique; a further advantage to his method is its ready implementation with non-linear least squares routines (with the obvious *caveat* that resulting distribution theory is unavailable).

PREDICTIONS

The data from 1947 to 1976 yielded an upper limit to Australia's population of 24.527 million, with a standard error of 3.098 million. This upper limit is almost eleven million higher than that predicted by the 1881-1933 data. This large difference in predicted saturation populations coincided with an increase in r, indicating different patterns of birth, death and migration in these two time periods, patterns that perhaps were also operational in the United States. On the basis of the 1947-1976 fitted logistic curve, a population of 20.381 million is predicted for the year 2011, with a standard error of 1.131 million. This estimate is close to the projection of 19.581 million made for that year by the Australian Bureau of Statistics and based on the component method.

The 1981 Australia census figure was 14.9233 million. The 1981 population predicted in 1980 by the Bureau based on the component method was 14.8137 million (-.73%) and their prediction made in 1978 was 14.674 millions (1977-78 Australian Year Book) with an error of -1.6%. The prediction from the logistic fit using information only until 1976 was 15.079 million (+1.05%).

Some further comparisons are possible. The Bureau's estimate of Australia's population as of December 1982 and released in 1983 (personal communication) was 15.2761 million, whereas the logistic prediction is 15.404 million, a difference of less than 1%.

Similar predictions may, of course, be made from the New Zealand logistic fit. For example, the post-war data yield an upper limit to New Zealand's population of 4.070 million. Perhaps subject to more immediate substantiation is the prediction from the post-war curve of a 1986 population of 3.3744 million, a figure slightly higher (+.2% to +.5%) than the various projections in the *New Zealand Official Yearbook.*

ACKNOWLEDGEMENTS

The authors are grateful to the Editor for his comments.

REFERENCES

Bacon, D. W. and Watts, D. G. (1971). Estimating the transition between two intersecting straight lines. *Biometrika* 58, 525-534.

Cameron, R. J. (Ed.) (1979). *Year Book Australia.* Australian Bureau of Statistics, Canberra, Australia.

Choy, J. H. C. and Broemeling, L. D. (1980). Some Bayesian inferences for a changing linear model. *Technometrics* 22, 71-78.

Department of Statistics (1975, 1984). *New Zealand Official Yearbook.* Government Printer, Wellington, New Zealand.

Hinkley, D. V. (1969). Inference about the intersection in two-phase regression. *Biometrika* 56, 495-504.

Keyfitz, N. (1968) *Introduction to the Mathematics of Populations.* Addison-Wesley, Reading, Massachusetts.

Leach, D. (1981). Re-evaluation of the logistic curve for human populations. *J. R. Statist. Soc. A* 144, 94-103.

Levenberg, K. (1944). A method for the solution of certain non-linear problems in least squares. *Quart. Appl. Math.* 2, 164-168.

Marquardt, D. W. (1963) An algorithm for least-squares estimation of non-linear parameters. *J. Soc. Indust. Appl. Math.* 11, 431-44.

Quandt, R. E. (1958). The estimation of the parameters of a linear regression system obeying two separate regimes. *J. Amer. Statist. Assoc.* 53, 873-880.

Quandt, R. E. (1960). Tests of the hypothesis that a linear regression system obeys two separate regimes. *J. Amer. Statist. Assoc.* 55, 324-330.

Verhulst, P. F. (1838). Notice sur la loi que la population suit dans son accroissement. *Correspondence Mathematique et Physique,* 10, 113-121.