MODELLING AND FORECASTING HUMAN POPULATIONS

USING SIGMOID MODELS

ROBERT RAESIDE

THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF Ph.D.

PRESENTED UNDER THE AUSPICES OF THE COUNCIL FOR NATIONAL ACADEMIC AWARDS

AND NAPIER COLLEGE

AUGUST 1987

Abstract

Robert Raeside

Modelling and Forecasting Human Populations using Sigmoid Models

Early this century "S-shaped" curves, sigmoids, gained popularity among demographers. However, by 1940, the approach had "fallen out of favour", being criticised for giving poor results and having no theoretical validity. It was also considered that models of total population were of little practical interest, the main forecasting procedure currently adopted being the bottom-up "cohort-component" method.

In the light of poor forecasting performance from component methods, a re-assessment is given in this thesis of the use of simple trend models. A suitable means of fitting these models to census data is developed, using a non-linear least squares algorithm based on minimisation of a proportionately weighted residual sum of squares. It is demonstrated that useful models can be obtained from which, by using a top-down methodology, component populations and vital components can be derived. When these models are recast in a recursive parameterisation, it is shown that forecasts can be obtained which, it is argued, are superior to existing official projections.

Regarding theoretical validity, it is argued that sigmoid models relate closely to Malthusian theory and give a mathematical statement of the demographic transition.

In order to judge the suitability of extrapolating from sigmoid models, a framework using Catastrophe Theory is developed. It is found that such a framework allows one qualitatively to model population changes resulting from subtle changes in influencing variables. The use of Catastrophe Theory has advantages over conventional demographic models as it allows a more holistic approach to population modelling.

Acknowledgements

I wish to express my gratitude to my supervisors:

Dr. D. Coleman, Department of Social and Administrative Studies, University of Oxford,

G.T. Fielding, Esq., Head of Department of Mathematics, Napier College, Edinburgh, and

D. Leach, Esq., Principal, Queen Margaret College, Edinburgh,

for all their helpful advice and encouragement.

I am indebted to Napier College and Lothian Regional Council for providing funding for this research. I wish to thank the staff of the Department of Mathematics and the Computer Unit for all their help.

Thanks are also due to Dr. C. Glasbey, AFRC Unit of Statistics, University of Edinburgh, for providing me with the basis of the non-linear regression computer program that is used in this thesis.

There are many others who have given assistance and stimulation, and this I acknowledge and appreciate. However, regarding the information presented in this thesis, I am entirely responsible for any errors, and any such deficiencies bear no reflection on those I have acknowledged.

Lastly, I am immensely grateful to Mrs. G. Gregory and Mrs. W. Wallace for their patience and ability in typing this thesis. Also Mrs M Cook for her great help at the end.

R. Raeside

Declaration

I hereby declare that:

All the work presented in this thesis has been carried out by myself and no part of this work has been submitted in support of another degree.

Robert Raeside

CONTENTS

Chapter 1 - Introduction - The Use of Sigmoids for Modelling and Forecasting Human Populations

1.1 Introduction

In this introduction the aim of this thesis is stated and its historical context defined. The specific areas to be explored are listed, followed by a description of how the thesis is organised in order to meet these objectives.

1.2 The Aim of the Thesis

The aim of the thesis is to investigate the use of models as a useful and credible means of modelling and forecasting human populations and component populations, concentrating attention particularly on sigmoids and other trend curves. The class of mathematical models represented by sigmoids are characterised by having an "S"-shaped graph with at least one point of inflection. Two of the simplest forms are the logistic model and the Gompertz model which are displayed in Figure 1.2.1.

The logistic model has the equation:

$$P_t = \frac{k}{1 - e^{\alpha-\beta t}}$$

which is a solution of the differential equation:

$$\frac{dP}{dt} = aP\left(\frac{k - P}{k}\right).$$

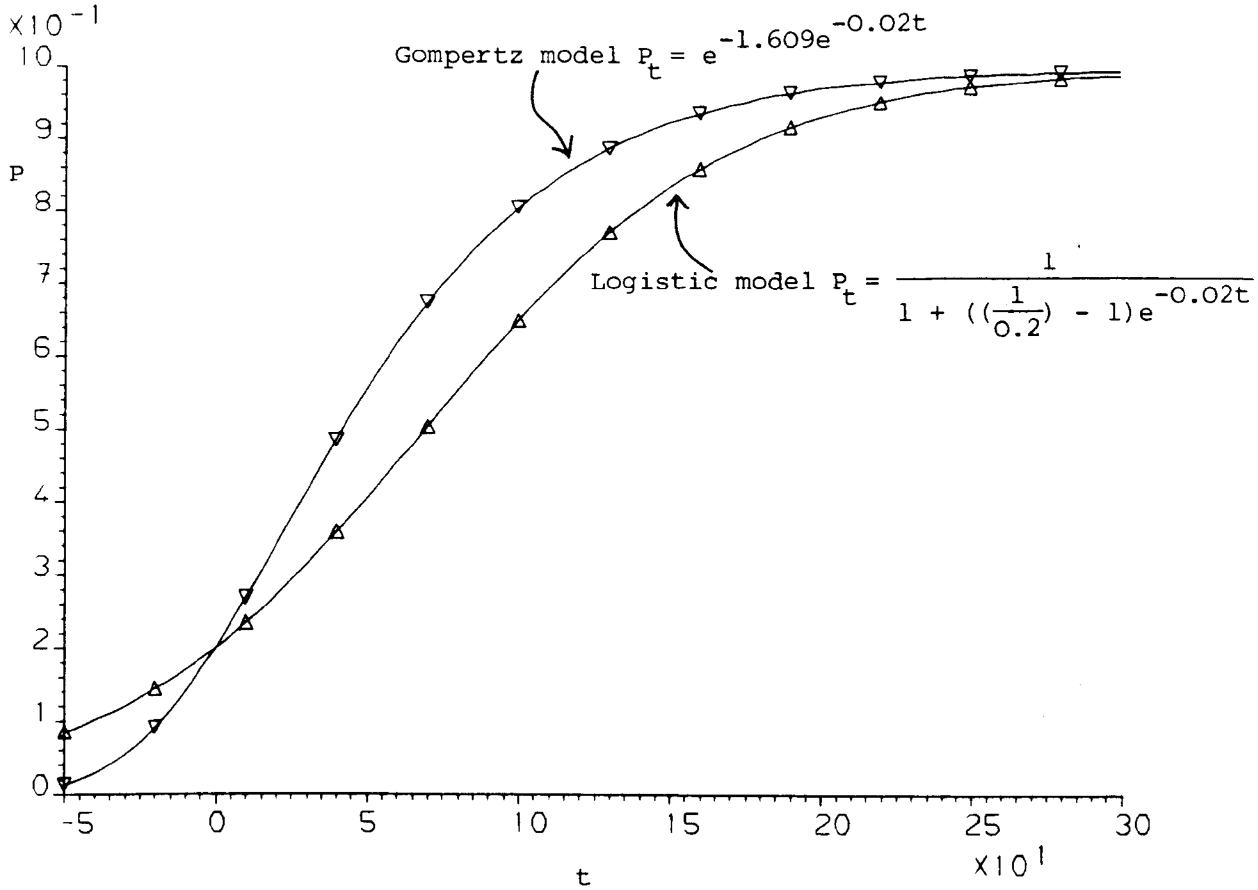
The Gompertz model has the equation $P_t = ke^{-e^{a-bt}}$ which can be derived from the differential equation $\frac{dP}{dt} = bPe^{a-bt}$.

Interest in these models arises due to their differential equations describing a process of accelerating growth followed by slowing and eventual cessation of growth.

The application of such models in demography are appealing to practical users if:

(i) they are simple to apply;

Figure 1.2.1: The three-parameter logistic and Gompertz models



(ii) assumptions can be clearly identified;

(iii) results of interpolation and forecasts are of acceptable quality;

(iv) the domain of the models' suitability can be established, and hence a criteria for testing the model is required;

(v) they can be applied to different populations at different time periods.

The above underpin the investigations carried out in this thesis.

1.3 Historical Development of Sigmoids

The first recorded use of an asymptotic model in demography is by Quetelet (1835) who proposed the use of an inverse square model to give an anti-Malthusian explanation of the growth of the Belgian population.

In 1838 Verhulst, a former student of Quetelet, argued that the checks on the growth of population increase in proportion to the population size. The S-shaped curve that Verhulst obtained he called the logistic model. However, due to the lack of census data and changes in society brought about by the onset of the Industrial Revolution which allowed apparently unbounded population growth (in the intermediate period), interest in modelling population waned and Verhulst's work was forgotten, according to Miner (1933). (Schitckzelle (1981) discusses further, in contemporary language, Verhulst's development of the logistic model and its application to Belgium in 1844).

Lloyd (1967) shows that the logistic model was rediscovered on at least five occasions before it was popularised by Pearl and Reed in the 1920s. These rediscoveries were in fields other than demography such as Robertson, who applied the logistic model to the growth of the individual organism and in 1915 compiled tables for the computation of Curves of Autocatalysis which were based on the logistic equation. Other rediscoveries were in modelling the growth of micro-organisms, plants, and the rate of chemical reactions.

Pearl and Reed (1920 and 1923) and Pearl (1925 and 1930) found that, on fitting the logistic model parameterised as:

$$P_t = d + \frac{k}{1 + e^{\alpha-\beta t}}$$

to population censuses, good descriptions of population growth could be obtained. In 1923 Pearl and Reed advanced the model as a first approximation to the law of population growth. Pearl (1925 and 1930) found the model to describe well the populations of Sweden, USA, France, Scotland, Austria, Belgium, Denmark, England and Wales, Hungary, Germany, Japan, Italy, Norway, Serbia, Java, Philippine Islands, the world and his home city, Baltimore. Pearl also found the logistic model suitable for animal population growth and growth of individual organisms.

Pearl (1925) concluded that "despite all the complexities of human behaviour, social organisation, economic structure and political activity, seem to alter much less than would have been expected the results of the operation of those biological forces which basically determine the course of populations of men as well as those of yeast cells".

Pearl (1925) goes further by writing: "In the matter of population growth there not only ought to be a law but six years of research has plainly shown that there is one".

Yule (1925) adds support to Pearl and Reed's work by making the use of logisitic models in demography the subject of his 1924 Presidential Address to the Royal Statistical Society. Yule illustrates the fitting of the model to the populations of England and Wales, France and the USA, and considers the consequences for the components of popoulation change (births, deaths and migration) for populations following such a curve.

In his work Pearl discovered that for some populations growth resumed (e.g. Germany and Japan). This meant that the simple four parameter model was no longer a good descriptor. Pearl and Reed (1923) advocated the use of the general model:

$$P_t = d + \frac{k}{1 + e^{a_1 t + a_2 t^2 + a_3 t^3 + \ldots + a_n t^n}}$$

After further empirical findings, Pearl (1924 and 1930) suggests that the human population grows in phases or cycles. In each phase, which Pearl refers to as an epoch, the population grows to a carrying capacity prescribed to the existing technology and social organisation.

Renewed growth (or decline) reflects changes in the carrying capacity which can arise as a result of new discoveries, better production techniques, change in land area, war or environmental changes. Thus Pearl views the history of human population growth as a series of epochs; in each epoch the growth can be described by a logistic model. Rather than use the general model above, Pearl found it better to use models for each epoch added in an escalated way. Figure 1.3.1 illustrates:

$$P_t = \begin{cases} d_1 + \dfrac{k_1}{1 + e^{\alpha_1 - \beta_1 t}} & \text{for } t_0 \leqslant t < t_1 \\[2ex] d_2 + \dfrac{k_1}{1 + e^{\alpha_2 - \beta_2 t}} & \text{for } t_1 \leqslant t < t_2 \end{cases}$$

In this thesis, the situation when there is a change in the pattern of growth which makes such modelling necessary is

Figure 1.3.1 - Escalated Sigmoids

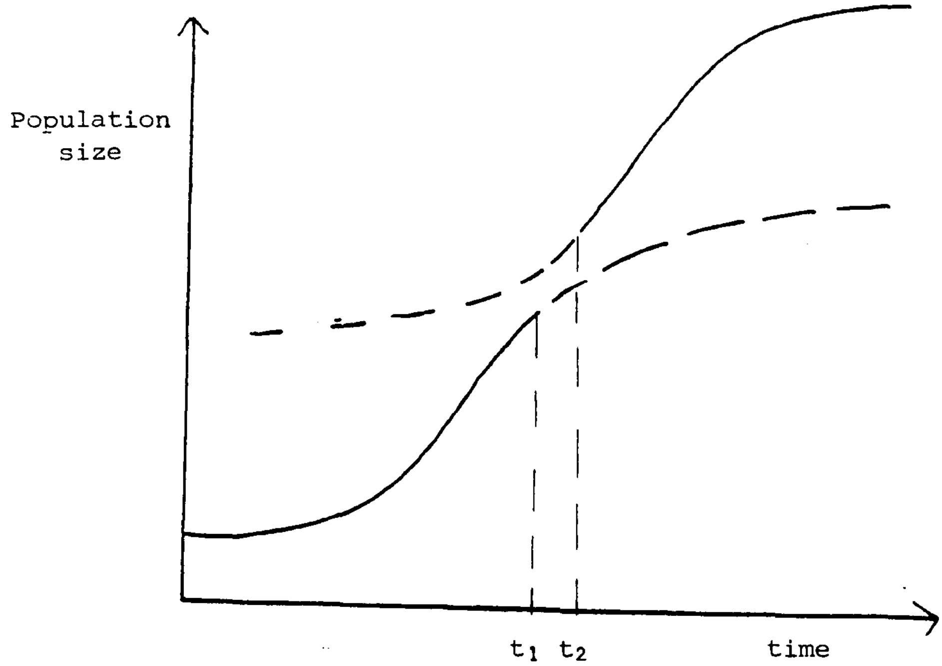


Figure 1.3.2a - Primative Cycle (Cycle I)

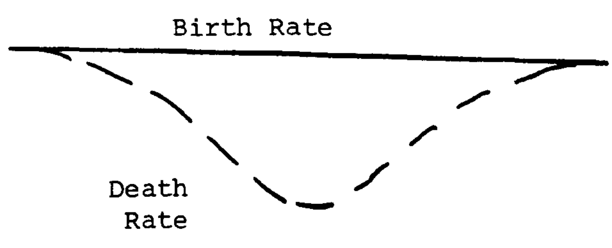


Figure 1.3.2b - Modern Cycle (Cycle II)

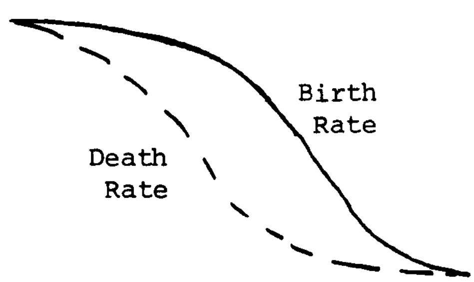


Figure 1.3.2 - Cowgill's Population Growth Cycles

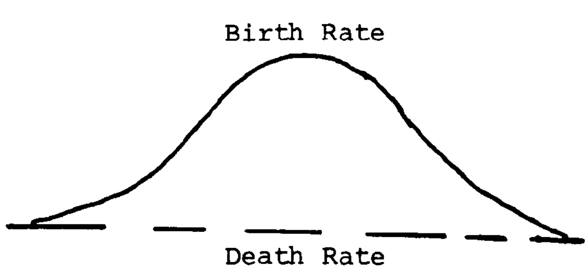


Figure 1.3.2c - Future Cycle (Cycle III)

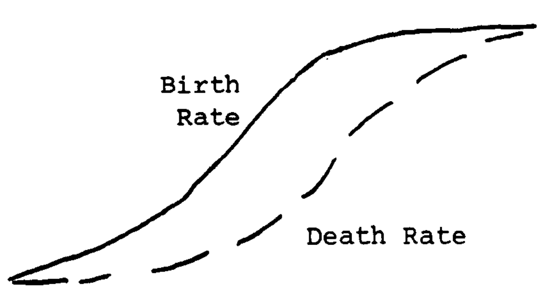


Figure 1.3.2d
Cowgill's Cycle IV

labelled a phase change. Cameron (1973) records qualitatively the history of European populations, suggesting that there are epochs or cycles of growth which can be represented by logistic models. Cowgill (1949) labelled these cycles as primitive, modern, and future, in which growth occurs as a result of different schedules of mortality and fertility. (An intermediate cycle resulting from rising birth and death rates is also identified.) These are illustrated in Figure 1.3.2.

Pearl and Reed (1920) and Pearl (1925) found extrapolations from the logistic model to be good forecasts for up to twenty years into the future, but, due to the possibility of phase changes arising, they were sceptical regarding the quality of longer lead time predictions, and hence caution must be attached to use of the upper asymptote.

Another sigmoid model - the Gompertz model - which was invented by Gompertz (1825) for the purpose of graduating mortality, has undergone a similar development as a growth model and is documented by Winsor (1932) who indicates that most of its applications have been to the biological growth of individual organisms.

The use of sigmoids for modelling growth gained popularity in many fields between 1920 and 1940. Hart (1945) lists over 100 applications of sigmoids to modelling social time series and Hart is led to the belief that the logistic model is a mathematical statement of a general law of growth. D'Arcy Thompson (1942) catalogues many examples of physiological growth that have occurred in a logistic fashion.

After 1930 interest in using sigmoids waned and the component method became popular for interpolating and forecasting populations replacing trend models. The main reason as to why sigmoids fell out of favour can be listed as:

(i) The major demographic interest to planners are component population, such as birth rates, and age structure models and forecasts, which did not appear to be readily available from trend models of the type considered (even though Lotka (1931, 1934) did show the relationship of births to a logistic model of total population). Dorn (1950) writes that the components are not available separately but the model "conceals their separate effects within its parameters. For the analysis of current demographic changes this is a serious defect".

(ii) Demographers were suspicious of the validity of sigmoid models. Wolfe (1927) launches a strong critical attack on Pearl's writings, being particularly critical of Pearl generalising support from controlled animal experiments. Dismissal of Pearl's work was probably encouraged by Pearl allying himself with the eugenics movement in America. Thus Wilson and Puffer (1933) commented: "If ... the statement that the logistic ... affords a rational law, to such an extent as to permit the extrapolation of the curve for forecasting purposes and the interpretation of the constants as constants of nature, we are forced to take serious exception to it ..."

(iii) The future ability of the logistic to describe the USA population on which Pearl based a great deal of his arguments became doubtful. This was particularly the case as the upper asymptote of 197 million was judged too low (this level was surpassed by 1970).

(iv) Even simpler methods fitted as well; Bowley (1924) comments: "For England and for France I find that an ordinary quadratic parabola (fitted by the method of least squares) gives figures almost identical with the logistic curve, for France even rather closer figures". Bowley goes on to state: "Thus on the test of goodness of fit alone the logistic equation has no special claims, so far as representation of past records is concerned".

(v) Demographic accounting methods (i.e. the component method which will be described in Chapter 2) had the illusion of greater accuracy (Ascher, 1979), and proved more information directly.

(vi) The idea of a law of growth is hard to prove conclusively and, in trying to do so, researchers have discredited the use of the model (this will be discussed further in Chapter 5).

(vii) In the earlier decades of this century, without electronic computers, the sigmoid models were awkward and time-consuming to fit.

(viii) There are statistical difficulties, such as the fact that population must be past the point of inflection before future growth can be described with even rough accuracy.

Feller (1940), when investigating the logistic model used by Thornton (1922) to model bacterial growth, and Reed and Holland

(1919) to model sunflower height, found other models to be as good descriptors of the data. Hence Feller argues that the fits of logistics are not remarkable and cannot be treated as evidence of a theory of growth. This raises a fundamental question which is addressed in this thesis.

But alternative modelling and forecasting schemes have not proved very successful. Particularly poor have been attempts at forecasting fertility (see Murphy (1980) and Ahlberg (1982)). Ascher (1979) also shows that component models perform no better (more often worse) than simple trend curves. However, the component models are logically consistent and would be good if fertility could be better predicted. Also modelling using sigmoids uses a more ambitious framework which incorporates feedbacks from the environment.

This has led to a revitalisation of interest in sigmoid models for human populations, examples of recent work are that of Stone (1979), Leach (1981), Oliver (1982), Tashchian and Tashchian (1984), and the work of Murphy (1980 and 1984). Jensen (1975) applies the logistic equation to fishery management to calculate yields from fish populations.

The use of sigmoid models has always been of interest in modelling and forecasting economic growth series such as sales of consumer durables and productivity. Recent examples of sigmoid applications in areas other than demography are to be found in the literature of Mar Molinero (1984), Marchetti (1985), Leach (1985), Leach and Wagstaff (1986), and Meade (1984 and 1985). Sigmoids have also been applied to the diffusion of innovation, see Metcalfe (1981) and van Wyk (1985).

1.4 Areas of Investigation in the Thesis

To establish the use of sigmoids in demography for modelling and forecasting populations, the following areas need to be explored and developed:

(i) the usefulness of sigmoids to describe population and to identify circumstances where sigmoids are unsuitable;

(ii) a methodology for fitting the models to available data;

(iii) methods of identifying and adjusting for phase changes;

(iv) the quality of forecasts and the suitability of a scheme for generating suitable forecasts;

(v) the use of the model to obtain disaggregated forecasts of births, deaths, migration and age structure, and hence make the use of sigmoids more appealing to practical users of demographic forecasts;

(vi) the applicability of the methodology to areas lacking in historical data series;

(vii) the rationality of implying that a good fit of a sigmoid model is evidence of a law of human population growth;

(viii) the relationship between sigmoid models and current demographic theories;

(ix) the possibility of presenting a unified modelling and forecasting scheme that is valid and fits the requirements of practical users.

To carry out this exploration, the thesis is organised as described in the following Section.

1.5 Organisation of Thesis

In Chapter 2 mathematical techniques for recognising the suitability of sigmoids to describe national human populations and means of fitting such models are reviewed. Non-linear least squares methods are found to be appropriate and a suitable computer program is developed. Emphasis is placed upon exploring the data (chiefly census records) to let it "speak for itself" in model selection.

A variety of simple time series trend models are compared. These models are fitted to populations from different parts of the world.

The populations modelled are grouped according to mainly economic status into developed, emerging and underdeveloped countries. Also considered are Island populations.

The problem of identifying and dealing with escalation (or phase changes) is given special consideration. The use of general type sigmoids is also examined.

Forecasting total populations is examined in Chapter 3. A number of different forecasting procedures are examined. These include straightforward extrapolation of models fitted to the data, ARIMA modelling and forecasting and forecasts made from recursive fits of the model by a simple approach based on only the most recent data points and by use of extended Kalman filters. Special attention is given to assessing the quality and validity of the forecasts generated. The mean absolute percentage error (M.A.P.E.) is used as an indicator of forecast quality. In this Chapter attempts are made to reflect the uncertainty associated with the forecasts. It is found that an approach based upon the use of the extended Kalman filter is the most suitable upon which to base a forecasting scheme.

In the next Chapter the problem of deriving component population models and forecasts is tackled. A methodology of obtaining stochastic models and forecasts (with confidence intervals) of births per year is presented, which is based on the work of Lotka (1931) combined with ARIMA modelling. Also modelled are changes in life expectancy at birth and the death rate, and exploration of the problems of modelling migration is made.

The "top down" models and forecasts of the vital rates of births and deaths are combined into a "bottom-up" component approach to model and forecast the sex-age structure of the population.

The age specific models are given high and low variants by supplying high and low confidence limits around births, and are amalgamated to act as a means of validating forecasts of total population made in Chapter 3.

In Chapters 3 and 4 the models are mainly applied to "data rich" developed nations. However, discussion is made of how the methods developed could be used for populations with sparse data.

In Chapter 5 the theoretical validity of using sigmoids is assessed with a view to dismissing the notions of the existence of a simple law of population growth. This is done by relating sigmoids to current demographic models and reviewing population change through history and in different cultural settings and how these changes have related to the sigmoid model. Also examined is the demographic interpretation of sigmoid parameters.

The modelling of population is pursued further in Chapter 6, to develop a holistic qualitative model of population change based upon the techniques of catastrophe theory. The model developed is used to give support to the use of sigmoids and to try to alleviate the remaining weakness by giving early indicators of the possibility of phase changes which require escalation.

Finally, in the last Chapter a summary of the main findings is presented and discussion of the practical use of the approach is made.

Each Chapter is subdivided into sections which deal with each major area covered in the Chapter.

In obtaining the results presented in this thesis, extensive use has been made of FORTRAN IV computer programs produced by the author which often utilise the NAG 10 and GINO-F subroutine libraries. The progam listings are available from the author. Also the computer packages MINITAB, GENSTAT and BATHSPLINE (Silverman and Watters, 1984) have been used extensively.

Chapter 2 - Modelling Populations

2.1 Introduction

In this Chapter the aim is to conduct a detailed investigation to determine the feasibility of devising a methodology for modelling populations using principally sigmoids and the presentation of results is the subject of this Chapter.

In Sections 2.2 and 2.3, the main trend models that can be used to model populations and a methodology for fitting these models to census counts are outlined. Methods of exploring the data that will both help to assess the quality of the data and to suggest appropriate models are discussed in Section 2.4.. In Section 2.5, the ideas developed regarding fitting are tested on simulated data and on some population data.

After reconsidering the fitting procedure in Section 2.6, the modelling is applied to actual population counts (usually census counts) in Section 2.7. This Section is split into seven subsections within which populations from the developed, emerging and less developed nations are considered. Also reviewed will be modelling the populations of relatively closed populations of Islands, and Quebec where, since 1870, migration has had little influence in shaping the population. Models of subpopulations, namely within the United Kingdom, and models of the reconstructed series given by Wrigley and Scofield (1981) are also considered in this Section.

In Section 2.8, more complex models which are still within a sigmoid classification are examined, in particular for cases where there have been changes in the phase of growth that require escalation or de-escalation.

2.2 Models of Population Growth

Willekens (1984) describes the main demographic modelling methods for populations. These are:

(a) Trend models - where the estimate of population (Z_t) is a function of time only. In general, $Z_t = f(t,\beta) + \epsilon_t$ where β is the vector of parameters in the model, ϵ_t is an error term.

(b) Time Series models - where the estimate of population at time t is a linear combination of previous observations. The model often has a form $Z_t = \Sigma \pi_j Z_{t-j} + \epsilon_t$, and such models are represented by autoregressive integrated moving average models (ARIMA, Box and Jenkins, 1970) which express the estimate Z_t as a linear combination of past observations and past errors. This is advantageous in allowing the model to be expressed by only a few parameters, a desirable principle, which Box and Jenkins (1970) refer to as parsimony.

(c) Component models - in component models the estimate of population, Z_t, is written as a function of the observation Z_{t-1} and on the components of change in one time period (usually one year). The components of change are births (fertility), mortality and migration. This model is represented by the balance equation (Keyfitz, 1968):

$$Z_t = Z_{t-1} + B - D + InM - OM$$

B = Births, D = Deaths, InM = Immigration, OM = Emigration, in one time period.

This "Bottom up" method relies on maintaining accurately and reliably large data sets.

In using component models for interpolation or extrapolation either the current component rates are assumed to be constant, which is not the case in the long term, or theoretically based assumptions made as to likely changes in the component rates, or other modelling methodologies are used to model the component rates.

For demographers and planners, the component approach, is generally the methodology used. This is because it is often the change in components and influences on the age structure of the populations that is required (total population size being of less interest).

(d) Explanatory models - such models express the observation Z_t as a function of variables that are thought to explain the variation in Z_t. Such models can be quantitative, such as Joshi and Overton (1984) or Ermisch (1982), when the model is written as a multiple regression model. Quantitative models have many difficulties that include problems with:

(i) quantifying "unquantifiables" such as attitudes;

(ii) uncertainties as to what are real influences and what are spurious arising out of interaction of variables;

(iii) determining the direction of causality;

(iv) in many cases empirical results are contrary to theory;

(v) the manner by which variables interact and their importance change with time;

(vi) interpretation of the models may be difficult. (Interpretation often relies on the presence of coefficients which are not statistically significant.)

(vii) models are non-parsimonious and not easily updated.

Alternatively, explanatory models may be qualitative, and both types will be discussed at greater length in Chapters 5 and 6.

The chief focus of this thesis is on the "Top Down approach" of modelling total population size and then obtaining components. Trend models and their time series representations (TSR) offer the most convenient methodologies.

The main trend models and their TSRs are now outlined (These are described in detail by Willekens (1984).)

1. Linear trend model

$$Z_t = a + bt + e_t \qquad \text{TSR} \Rightarrow Z_t = Z_{t-1} + b + \epsilon_t$$

Due to having a constant rate of growth (b), such a model is only likely to be applicable in the short term.

2. Polynomial models

$$Z_t = \sum_{i=0}^{n} a_i t^i \qquad \text{TSR} \Rightarrow Z_t = \sum_j \pi_j Z_{t-j} + \epsilon_t$$

(In this thesis polynomial models will be restricted to quadratics and cubics.) Due to their flexibility (especially cubics and higher orders), polynomial models are likely to be good for interpolation, but, since they tend to be unbounded outside the range of the data, they are unsuitable for long-term extrapolation. The quality of short-term extrapolations (as will be examined in Chapter 3) is related to the positions of stationary points in the models.

3. Exponential growth models

$$Z_t = ae^{rt} + e_t \qquad \text{TSR} \Rightarrow Z_t = e^r Z_{t-1} + \epsilon_t$$

r = the rate of growth which is constant.

This model can be modified by adding a constant (b) so that $Z_o = a + d$.

This model is used extensively in demography often to show the consequences of unchecked population growth. This predicts unbounded population growth and has been often incorrectly termed the Malthusian model (e.g. Burghes and Borrie, 1981). (The fallacy of this terminology will be discussed in Chapter 5.)

When r is less than zero, declining populations are modelled.

The perturbed model $Z_t = b + ae^{rt} + \epsilon_t$ is also of interest. When a and r are negative, this models population growth that monotonically tends to an upper asymptote (b). According to such a representation of the model, the difference between the maximum population and the actual population declines exponentially. Alternatively, when the parameters are positive, exponential growth is modelled.

4. Sigmoids

These are the main models examined in this thesis and exhibit the desirable properties of upper and lower asymptotes and have points of inflection marking changes from accelerating growth rates to declining growth rates.

4a. Symmetrical Sigmoids - these are models which are symmetrical about a point of inflection of which the logistic is an example.

The logistic model $Z_t = \dfrac{Z_\infty}{1 + (Z_\infty/Z_0 - 1)e^{-bt}} + \epsilon_t$

where Z_∞ is the upper asymptote and has TSR

$y_t = (1 + b)y_{t-1} + by_{t-2} + \epsilon_t$ where $y_t = 1/Z_t$. This is an ARIMA (2,1,0) model in the Box-Jenkins framework.

The differential equation of this model is:

$$dZ_t/dt = Z_t b(1 - Z_t/Z_\infty)$$

The point of inflexion in the curve of the model is $Z_\infty/2$.

A useful modification of this model is to add a constant (d),

$$\text{i.e. } Z_t = \frac{Z_\infty}{1 + (Z_\infty/Z_0 - 1)e^{-bt}} + d$$

with differential equation:

$$dZ_t/dt = b(Z_t - d)(1 - Z_t/Z_\infty)$$

and the upper asymptote and point of inflexion change to $Z_\infty + d$ and $Z_\infty/2 + d$ respectively.

These models have already been described in Chapter 1. There are several reparameterisations of this model of which

$$Z_t = \frac{Z_\infty}{1 + e^{a+bt}} = \frac{\alpha/\beta}{1 + e^{-\alpha(t-t_0)}}$$

is a form that appears in demographic literature (Pearl and Reed, 1920s, Willekens, 1984).

A similar model is the model based upon the hyperbolic tangent:

$$Z_t = \tfrac{1}{2}Z_\infty[1 + \tanh \tfrac{1}{2}bx].$$

This model has had little application in demography. (A notable exception is Lotka (1931), as will be discussed further in Chapter 4.)

4b. Asymmetric Sigmoids

The main asymmetric sigmoid considered in this thesis is the Gompertz Model:

$$Z_t = ae^{be^{-ct}} + \epsilon_t \quad \text{with TSR}$$

$$y_t = (1 + b)y_{t-1} - by_{t-2} + \epsilon_t, \text{ where } y_t = \ln Z_t$$

and this has a point of inflection at $Z_t = Z_\infty/e$.

The model can be made more useful by adding an offset, i.e.

$$Z_t = d + ae^{be^{-ct}} + \epsilon_t.$$

The three parameter logistic and Gompertz are compared in Figure 1.2.1 for growth between 0 and 1.

Other forms of sigmoids, including a generalised form of sigmoids that has as special cases exponential, geometric and logistic growth, are reviewed by Stone (1979). Some of these are considered in the context of population modelling in Section 2.8.

From this outline, one can conclude that as a class of models, for "Top Down" modelling, trend models are appropriate. Also it appears sigmoids, for theoretical reasons, are superior to other trend curves as they:

(a) have upper and lower limits;

(b) do not have fixed rates of growth;

(c) stem from differential equations that model growth process.

Empirical investigation will be made of the simple sigmoids discussed and their performance will be compared to other trend curves.

In order to do this, a means of fitting these models to data has to be developed and some explanation of how the models relate to the data to be used is required. This is the subject of the next four Sections.

2.3 Fitting Sigmoid Models to a Time Series of Population Counts

To fit non-linear trend models Z(t) that were outlined in the previous Section to a time series of n population counts P(t), various researchers, such as Nelder (1961), Bhattacharya (1966), Oliver (1964, 1966, 1969 and 1982) and Tashchian and Tashchian (1984), have suggested using optimisation methods based on "least squares" algorithms. Oliver (1982) advocates minimising the negative log likelihood function:

$$\ln \mathfrak{L} = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln \sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{n} e^2{}_t$$

where e_i = observed population - fitted population = $P_t - \hat{P}_t$.

Essentially parameter estimates are obtained by the method of minimisation in a non-linear least squares algorithm. New parameter estimates ($\underline{X}_k$) are given by:

$$\underline{X}_k = \underline{X}_o + \Delta\underline{X} = \underline{X}_o - [A^TA]^{-1}A^TF$$

where $F = \sum_{t=1}^{n} (P_t - \hat{P}_t)^2$ is minimised and A is the matrix of partial derivatives of the function (A^TA is the hessian matrix H).

For the trend models considered which have derivatives that are easily computed, then efficient algorithms to achieve this iterative minimisation are based on Newton's method with Marquandt's modification to the hessian matrix to allow variable step sizes and so hasten convergence (Schwefel, 1981, see also Adby and Dempster, 1974).

This method is used in the thesis in computer programs and when use is made of the OPTIMIZE directive in the GENSTAT computer packages and the non-linear FORTRAN NAG (Rev 10) library routines.

Estimates of the standard errors of the parameter estimates and correlation between the parameter estimates can be obtained from the asymptotic variance-covariance matrix whose inverse is computed by a scheme given by Oliver (1982). However, this methodology is only appropriate when assumptions that the residuals are serially independent and have constant variance are satisfied (Stone, 1979, Nelder, 1961).

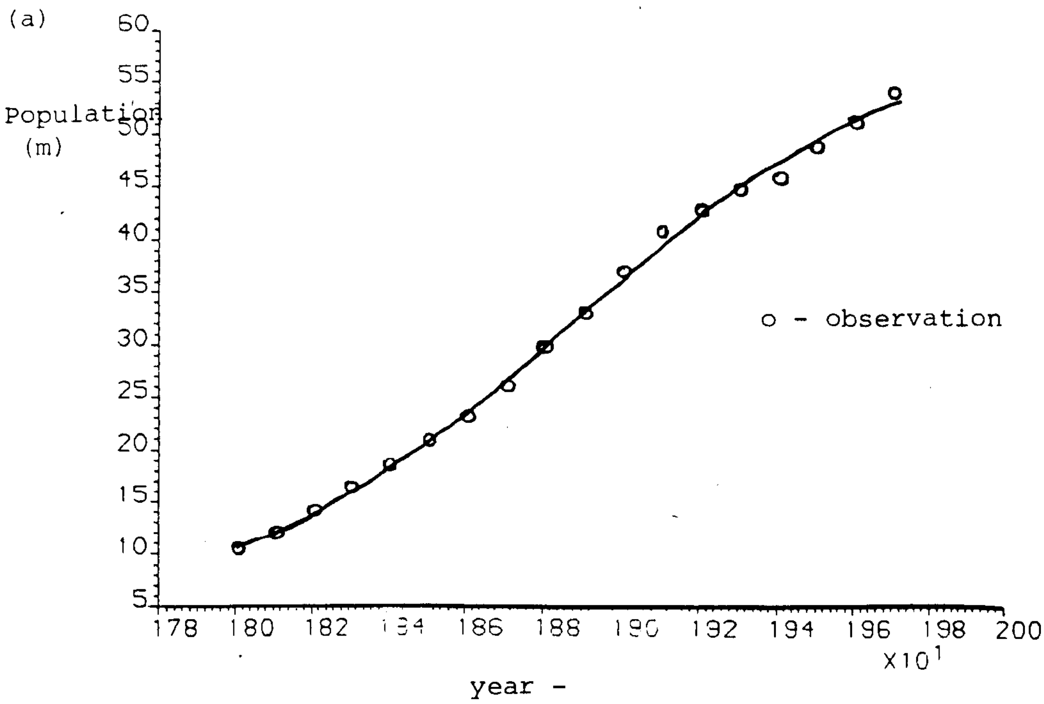
Figure 2.3.1, which compares a curve generated by a four parameter logistic (fitted by Oliver, 1982) to the population of Great Britain over the period 1801 to 1971, suggests the above assumptions are not the case. Observations drawn, particularly from Figure 2.3.1b, show patterning to the residuals suggesting serial correlation giving a cyclical pattern, with a period of between thirty and fifty years, in the data. This serial correlation which, on reflection is to be expected, as:

- The population size at the next census depends on the population alive at the current census - most demographic series are inherently correlated.
- There are likely to be cyclical variations in fertility (and migration) in response to cycles of economic growth and perhaps by changes in cohort sizes (Easterlin, 1968, 1978, 1980) - this will be considered further in Chapter 5.
- The effects of errors in enumeration could be reflected in a cyclical pattern.
- Influences of changes in age structure.

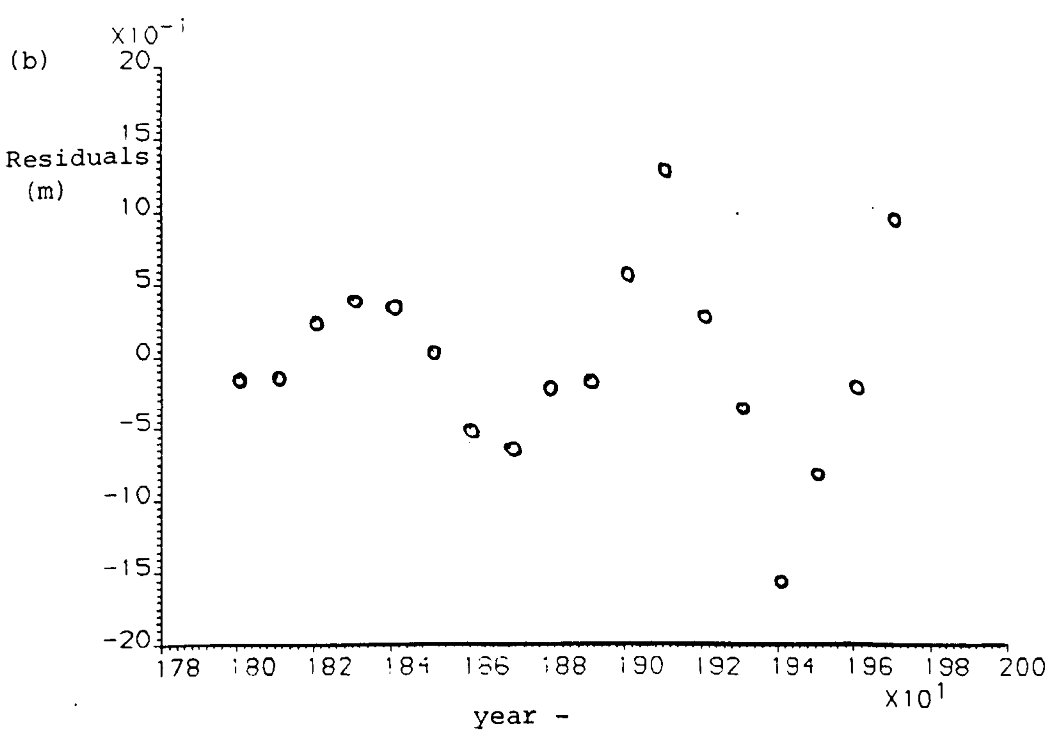
Sandland and McGilchrist (1979) state that the effect of this strong dependent structure in the growth process is to make fitting by ordinary least squares inefficient in that, unless close initial approximations are made, it can be difficult to get the least squares routine to converge. Oliver (1981) argues that this residual autocorrelation can be safely ignored and the

Figure 2.3.1: Oliver's (1982) 4-parameter logistic model of Great Britain

(a)



(b)



effect of this "would be to cause underestimation of the sampling variability of the parameter estimates". This is examined further in Section 2.5.

Various other researchers have considered the problems of correlated errors, such as Vieira and Hoffman (1977) who consider an additive and a multiplicative error structure, and Sandland and McGilchrist (1979) who point out that the shortness of the series considered precludes the use of complex ARIMA processes for errors and so the incorporation of an ARIMA error term must be restricted to simple parsimonious cases.

Glasbey (1979, 1980) applies a standard maximum likelihood method to optimise a function with an autoregressive error structure incorporated into it. Glasbey's modelling will now be summarised.

The form of model fitted is $P_t = \hat{P}_t + \sum_{t=1}^{n} \alpha_t e_{n-t} + \gamma$ which can be thought of as a model with a deterministic part and a stochastic part. Incorporating this into the negative log likelihood function gives:

$$\ln \mathfrak{L} = \frac{n}{2} \ln \sigma^2 + \frac{(n-1)}{2}\left\{\ln(1-\rho^2) + \frac{A + (1+\rho^2)B - 2\rho C}{2\sigma^2(1-\rho^2)}\right\}$$

where $A = e_1{}^2 + e_n{}^2$, $B = \sum_{t=2}^{n-1} e_t{}^2$, $C = \sum_{t=1}^{n-1} e_t - e_{t-1}$, ρ is the first order autocorrelation coefficient. This models for an error term which is normally distributed with a zero mean and variance $\sigma^2\rho$.

If ζ_0 is the current parameter estimates, then, on each iteration, ζ_0 is adjusted to $\zeta_0 + \sigma\zeta$ where:

$$\sigma\zeta = -H(\zeta_0)^{-1} - (\sigma\mathfrak{L}/\sigma\zeta_i/\zeta_0).$$

H is a positive definite matrix of partial derivatives. In algorithms used in this thesis, iterations are continued until the largest change in a parameter estimate is less than 10^{-6}.

An estimate of the covariance matrix is given by the inverse of the hessian matrix. The variance (σ_o^2) and autocorrelation (ρ_k) are estimated by:

$$\sigma_o^2 = (\sum_{t=1}^{n} e_t^2)/n \qquad \rho_k = \sum_{t=1}^{n} \frac{(e_t - \bar{e})(e_{t+k} - \bar{e})}{\sum_{t=1}^{n} (e_t - \bar{e})^2}$$

The autoregressive model of the errors is directly attributable to a Box and Jenkins (1970) autoregressive series. The procedure for determining the autoregressive terms is described by Glasbey (1980). Fitting such a many parametered model often proves to be numerically problematic, so the model's fit is optimised in a nested manner in that first the deterministic part is fitted, and then the errors are modelled. This is carried out iteratively until a suitably low and suitably distributed residual is obtained. (This method was suggested by Ross (1970).) A computer program, PTRENDS is used to carry out this modelling, which is a modification of a routine supplied by Glasbey and is described by Raeside (1984).

Oliver (1981) considers the estimate of the parameters to be asymptotically independent of the estimate of the autocorrelation and so considers the above nested approach to be feasible. Glasbey (1979) confirms the asymptotic independence of the autocorrelation and the parameter estimates, but he also shows that in small samples the estimates of the parameters of the sigmoid function are influenced by the presence of serial correlation in the residuals. Mar Molinero (1984) points out that, because of the complicated interactions involved, it is not clear in what way the estimates of the parameters or the estimates of their standard errors will be affected if the presence of residual autocorrelation is ignored.

Thus, methodology developed by Glasbey (1979) would appear to be the most suitable to use to fit non-linear regression models.

Mar Molinero (1984) also suggested incorporating other independent variables into the model to get an uncorrelated error. However, this is unsatisfactory for a simple modelling procedure as it is by no means an easy task to decide what variables other than time should be included (a point which will be taken up further in Chapters 5 and 6).

From Figure 2.3.1b another deficiency is apparent with ordinary least squares as carried out by Oliver (1982). There seems to be inhomonogeneity in the variance of residuals in that there is much larger variance, associated with more recent times, than at the beginning of the series. Possible remedial measures are to transform the data (Box and Cox, 1964). Logarithmic or reciprocal transformations are possibilities.

Weighting the data, perhaps, by changing the error term in the likelihood function to

$$e_t = \frac{P_t - \hat{P}_t}{\hat{P}_t}$$

may also stabilise the variance of the residuals. The use of weight has the appeal of using a proportionate error term which will allow comparisons across different populations and within a population at different times. However, it does have an undesirable property in that equal weight is attributed to each point in the series in the least squares minimisation. This is disadvantageous because, in ordinary non-linear least squares regression, the largest data points are attributed most weight. These are usually the most recent and reliable counts of the population in the data sets considered in this thesis.

In the fitting methodologies considered, initial estimates of the parameters are required. These can be obtained by trial and error (the modified Newton's method is usually efficient even for initial estimates that are far from the optimal values), or by solution of the p simultaneous equations (where p is the number of parameters). This is used in the computer program PTRENDS. (Some graphical means of obtaining initial estimates are commented on in the next Section.)

Before stating the fitting methodology, some exploratory procedures prior to fitting are considered in the next Section, and preliminary fits of USA, France, Sweden and simulated data are investigated in Section 2.5.

2.4 Exploring the Data

Before embarking on fitting the models to population it is useful to explore the data. Exploratory data analysis in allowing the data to "speak for itself" is useful to:

(i) determine the appropriate parametric model;

(ii) initially set the magnitude of parameters;

(iii) detect odd points in the data which could be classed as outliers - where perhaps there was an error in the census enumeration or additional land and people were either gained or lost from a nation;

(iv) detect changes in growth rates that might indicate changes in the "phase" of population growth that would require escalated or de-escalated modelling.

There are a variety of ways to explore the data; the methods that are used in this thesis are:

(a) Plots of the data - plotting the data against time is often a useful and simple way of choosing the parametric model, estimating N_0 and possibly N_∞ if a sigmoid is suitable, and detecting points that are "out of trend" with the bulk of the data. Other useful plots are plotting first differences of the population counts against either time or population size. If a sigmoid is to give a good model of the data, then this plot should be of inverted 'U' shape and is symmetrical if a three parameter logistic gives a good model. The peak of the inverted 'U' is the point of inflection and is at half the upper asymptote. A plot worth considering is of first differences divided by the population size against population size. If a three parameter logistic is a good model then this plot should form a straight line with a negative slope (see Hotelling, 1927). If this line is extrapolated to the horizontal axis, one obtains the point where $N_t - N_{t-1} = 0$, i.e. the upper asymptote. (If a Gompertz model is more suitable, then this plot would appear as a convex curve.)

(b) Non-parametric Modelling - the use of non-parametric modelling as a means of modelling data (y_t) and allowing the data to suggest a suitable parametric form is advocated by Silverman (1985) in a paper which received a great deal of critical acclaim. Silverman develops his approach out of the classic least squares method of minimising the residual sum of squares (RSS), $\Sigma(y_i - g(t_i))^2$ where g is the estimated curve. The RSS can be reduced to zero by choosing g so that it interpolates the data (which is distinct in cases considered). However, rapid local fluctuations are considered implausible and hence an undesirable feature of g. To overcome this, Silverman adds a roughness penalty into the computation of the RSS to represent the cost of rapid local variations. This gives the modified residual sum of squares term to be:

$$S(g) = \Sigma\{y_i - g(t_i)\}^2 + \alpha\int g''(X)^2\, dx$$

where $\int g''(X)^2\, dx$ represents the roughness penalty and α is a smoothing parameter that determines the "trade-off" between a good fit and excessive rapid local variation.

Silverman points out that the solution curve $\hat{g}$ has the following properties which are not imposed but arise automatically from setting the roughness penalty to $\int g''(x)^2\, dx$:

(i) it is a cubic polynomial between adjacent data points;

(ii) the curve has continuous first and second derivatives at the data points;

(iii) the second derivative is zero, i.e. g is linear, outside the range of the data.

The curve which satisfies properties (i) and (ii) is called a cubic spline.

Silverman and Watters (1984) provide a computer package "BATHSPLINE" which fits such curves to data using the modified residual sum of squares, S(g) in the covergence criteria. This package was used in exploring the population data. The package gives:

(i) plots of the spline compared to the data and 95% confidence intervals for the fitted curve, which are found from a Bayesian formulation (see Silverman, 1985);

(ii) plots the derivative (and inference regions on the derivative) of the spline;

(iii) plots the residuals from the fit against time (these can can be raw, standardised or absolute residuals).

Silverman recommends that the choice of the smoothing parameter is made by "trial and error" (i.e. pick a value of α and examine the plot and its residuals); alternatively α can be chosen automatically in the package by using a cross-validation procedure, described by Silverman (1985).

The procedure of using splines is useful in suggesting a parametric form; as an example Silverman (1985) gives the study of growth curves.

Using this type of spline analysis is of importance in this thesis for the following reasons:

(i) To give a benchmark for comparing parametric models, the spline being viewed as the best possible description of the data. A parametric model whose estimates fall within the 95% confidence intervals on the spline is judged a good model.

(ii) To detect outlying data points that require specific investigation.

(iii) To determine the maximum rate of growth from the plot of the derivative of the spline. This is useful in giving an initial estimate of the upper asymptote of sigmoids used in modelling.

(iv) An inverted 'U' shape of the plot of the first derivative of the spline indicates the appropriateness of sigmoids as parametric models.

(v) Examination of the first derivative of the spline is also useful in determining changes in "phases" of growth. Where there is renewed growth there should be multiple peaks in the plot of the first derivative - this would indicate escalation.

(vi) Examination of the plots of the residuals from the fit of the spline to the data points will indicate if transformation or weighting of the data is required to give residuals which are independent, having zero mean and constant variance.

However, the methodology and package developed by Silverman assumes that the data is uncorrelated, yet the data considered in this thesis is correlated. A way of overcoming this is suggested by Diggle (1985) by replacing the RSS component by an autoregressive formulation:

$$\mathrm{RSS}\ (\alpha,\rho) = \sum_{i=2}^{n} [\{y_i - g(t_i)\} - \rho\{y_{i-1} - g(t_{i-1})\}]^2$$

where ρ is the autocorrelation coefficient.

Robinson (1985) points out that such an alteration has repercussions right through the methodology. Thus complexity of the method is increased.

Since this analysis is only to be used for exploratory purposes, modifications to allow for autocorrelation in data shall not be considered further other than the observation that the size of confidence intervals will increase (possibly double Glasbey, 1979). (Also Silverman does not consider that there is much distinction between the smoothing parameter and the serial autocorrelation coefficient.) Thus the package will be used in an unmodified form.

(c) Window analysis - this is an extension of a method used by Leach (1981). This analysis is performed by choosing a model and fitting it to a portion of the data (a window). This window is enlarged to include an additional data point and the model refitted. This is continued until all data points are included. At each iteration* the initial parameter estimates are taken from the previous fit. This is also carried out backwards starting with the most recent portion of the data. A window of constant size (number of data points) can be moved through the data string as well. On completion of the iterations* the values of the parameters and the residual sum of squares (or likelihood) are plotted against iteration number. By examining variation in the parameters and residual sum of squares, it is possible to infer that a discontinuity is encountered where there are sudden changes in the magnitude of the parameter estimates or rapid rises in the value of the residual sum of squares. A discontinuity could result from a change in the phase of growth requiring escalation or de-escalation, an outlying point or the model being of an inappropriate form. Leach (1981) uses the forward moving window (considering the root mean sum of squares of the residuals (r.m.s.s.) rather than just RSS). He concludes that, after the point of inflection in 1901 there is stability in the upper asymptote and the r.m.s.s. for the population of Great Britain, but this is not the case for the population of the USA where

* Iteration, is taken to mean a recursive movement of the window.

variation in parameters and r.m.s.s. suggests the need for escalation after 1950. Leach, using this method, also detects a requirement for de-escalation in modelling the population of Scotland after 1921.

Leach (1981) points out that the parameters of the logistic are inter-related. This is confirmed by Oliver (1982). The estimate of the C parameter is highly negatively correlated with N_∞, and N_0 is positively correlated with N_∞. Thus, in carrying out the window analysis, only movement in the upper asymptote and the RSS (or the negative log likelihood) will be considered since variations in other parameters will be related to movements of N_∞. This window analysis will be mainly carried out with unweighted data as variations should be greater and hence more detectable.

(d) Linear time series models - the time series representations of the simple trend models, which were formulated by Willekens and stated in Section 2.2, can be easily fitted before going on to carry out non-linear modelling. A simple regression of $1/N_t$ on $1/N_{t-1}$ and $1/N_{t-2}$, for example, is a useful way to give a quick check to test if the three parameter logistic model is appropriate. If a three parameter logistic model is suitable, then the form of the regression equation should be:

$$1/N_t = (1 + b)1/N_{t-1} - (b)1/N_{t-2}$$

and residuals should be random with zero mean and constant variance.

This can easily be examined using, for instance, the "MINITAB" computer package. McNeil (1974), working within an exploratory data analysis framework, uses such regressions to obtain suitable transformations and model parameterisations to overcome curvature and correlations in the data and give residuals which have zero mean and are randomly scattered. He demonstrates his modelling on the USA population. McNeil's methods are time-consuming when compared to other ways of exploring the data and this method is only suitable for equally spaced data, so this method will not be considered as a means of exploring the data.

The simpler technique above is only convenient for situations where the data is equally spaced, and so such analysis is only adopted when the census counts are regularly spaced.

(e) Cumulative sum (cusum) plots - Sandland and McGilchrist (1979) use these plots (as well as plots of relative average growth rate against time) as an aid in the detection of changes of growth phases. The cusum is formed by plotting $\sum_{j=1}^{N} (R(j) - \bar{R})$ against time where R(j) is the relative growth rate $\left(\frac{N_t - N_{t-1}}{t_2 - t_1}\right)$, $\bar{R}$ is given by $\sum_{j=1}^{N} R(j)/n$. Sharp peaks in the cusum plot indicate changes of phase.

2.5 Simulated Logistic Populations and Investigation of Fitting Methods

Simple simulations were carried out to produce populations that have grown in a logistic manner in order that the fitting methodology can be investigated and results compared to known values. The simulations did not consider the complexity of age structure and assumed that the probability of reproducing or dying is independent of age. The simulation methodology adopted was that described by Pielou (1977) and is outlined in the following algorithm:

(i) Assume that the birth rate decreases linearly and the death rate increases linearly as the population grows. Pielou shows that under this assumption the probability of a birth $Pr(N \to N+1) = (a_1N-b_1N^2)/((a_1+a_2)N-(b_1-b_2)N^2)$ and the probability of a death $Pr(N \to N-1) = (a_2N+b_2N^2)/((a_1+a_2)N-(b_1-b_2)N^2)$ where N is population size and $a_1+a_2 = c$, the growth parameter, and $(a_1+a_2)/(b_1+b_2) = N_\infty$, the upper asymptote.

(ii) Set the initial population size N at time t.

(iii) Generate a uniform pseudo-random number R from the interval [0,1].

(iv) If $R \leq Pr(N \to N+1)$ then let $N = N+1$, otherwise let $N = N-1$.

(v) Calculate the time interval Δt between events, where $\Delta t = ((a_1+a_2)N - (b_1-b_2)N^2)^{-1} \ln(1-R)$, and update time, i.e. $t = t+\Delta t$.

(vi) Repeat steps (iii), (iv) and (v) until either t reaches an upper level t_{max} or N reaches an upper level N_{max}.

In this simulation the uniform pseudo-random number generator, AS183, devised by Wichmann and Hill (1985) was used.

Although such a simulation cannot be regarded as representative of human population growth, the results will serve to examine and test the fitting algorithms that are considered.

Simulations were run to give populations with upper asymptotes 100, 1000, 10,000 and 100,000. Some of the typical simulations obtained are graphically illustrated in Figure 2.5.1 which shows the stochastic growth of the population compared to the smooth path of the deterministic model.

Examination of the graphs in Figure 2.5.1 indicates three important features:

(i) There is closer correspondence to the deterministic path with large populations.

(ii) During the growth part of the curve the stochastic path "wanders" away from the deterministic path to remain either above or below the deterministic path until the upper asymptote is approached. (This behaviour is not unusual in stochastic processes, see Feller, 1950).

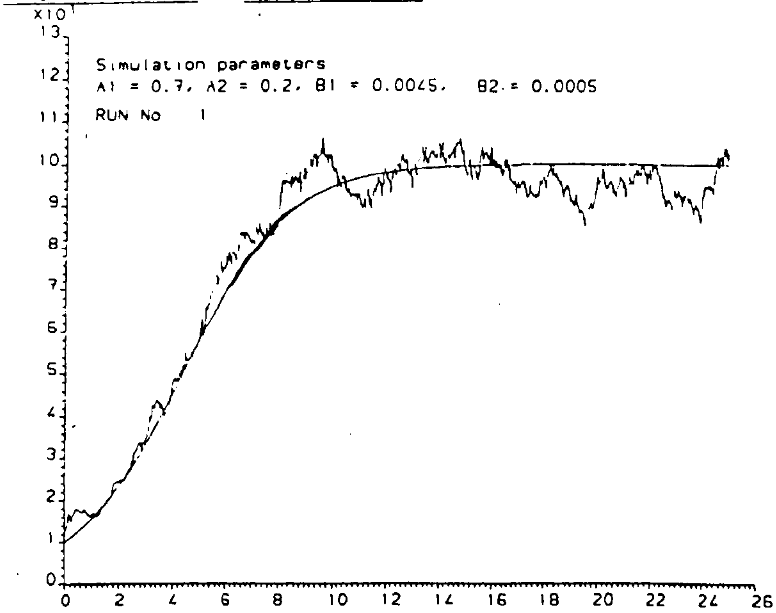
(iii) On nearing the upper asymptote the stochastic path takes a wave form oscillating around the upper level with a relatively long period.

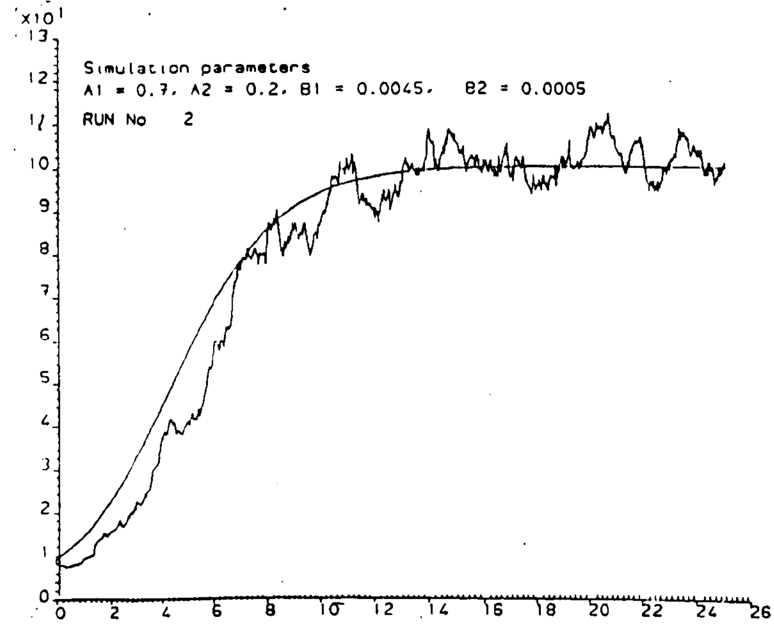
The effect of introducing age structure into the simulation is to superimpose a wave of variable period on the stochastic part (see Ph.D. thesis of Norden, 1980).

The simulations were used to assess the ability of the fitting algorithms developed and to investigate the effect of using Glasbey's (1980) method to account for the inherent correlations in such time series.

Table 2.5.1 gives the results of comparisons of the estimates of parameters, obtained from fits of the three parameter logistic

Figure 2.5.1: Simulations





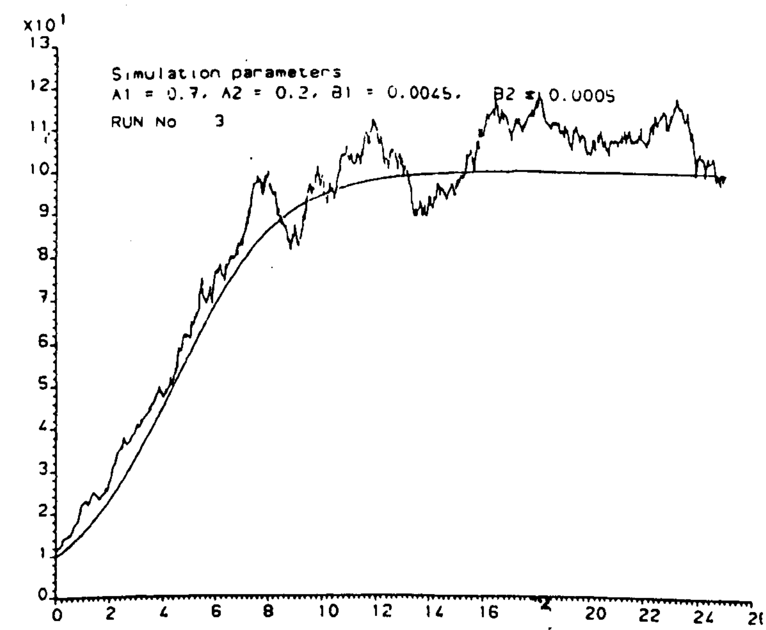
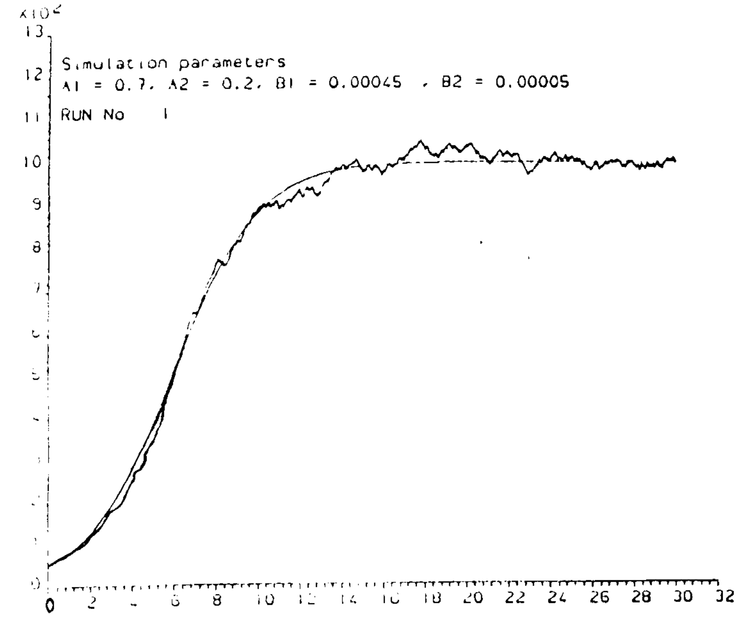
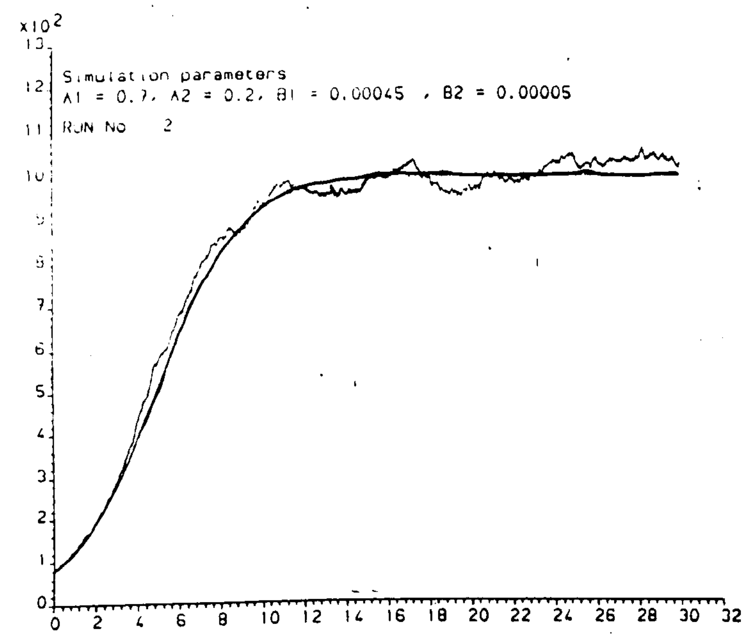


Figure 2.5.1 (continued)





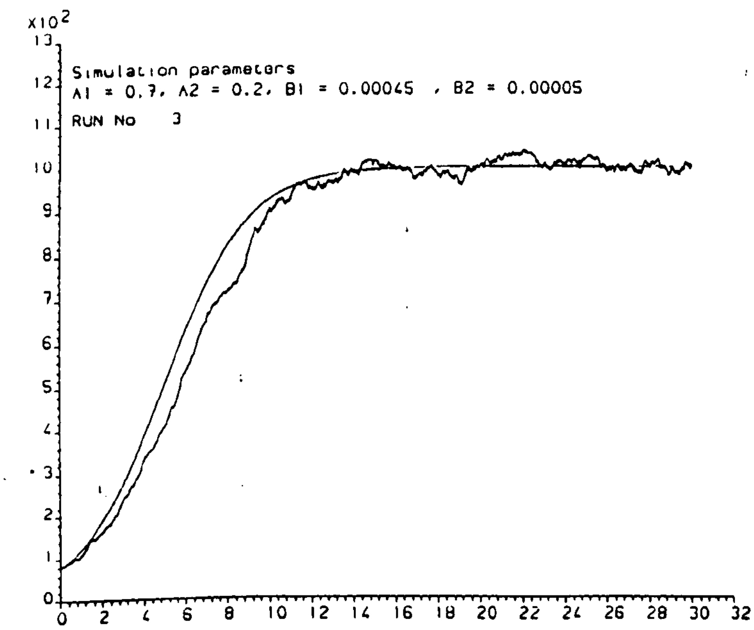
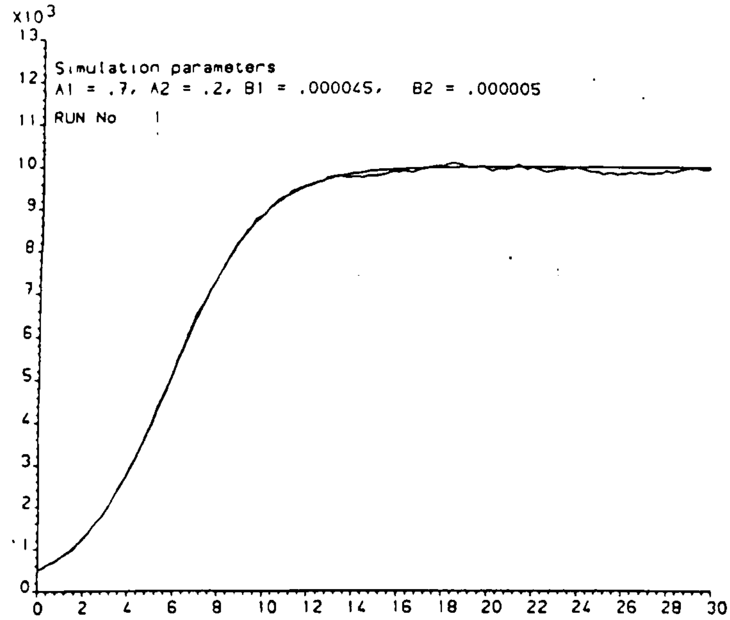
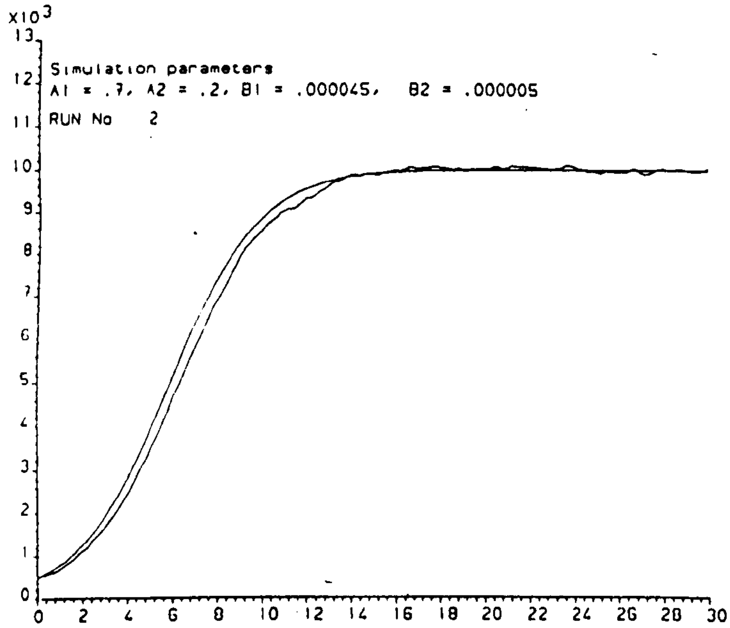


Figure 2.5.1 (continued)





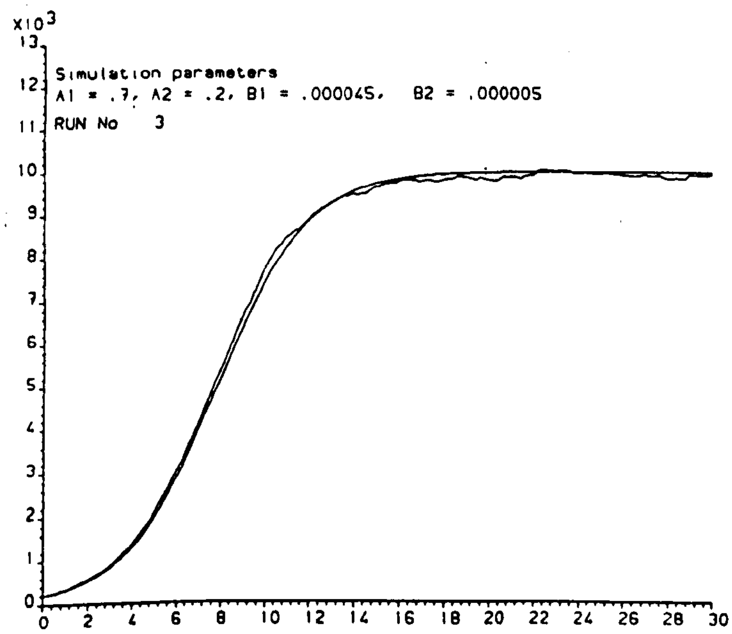
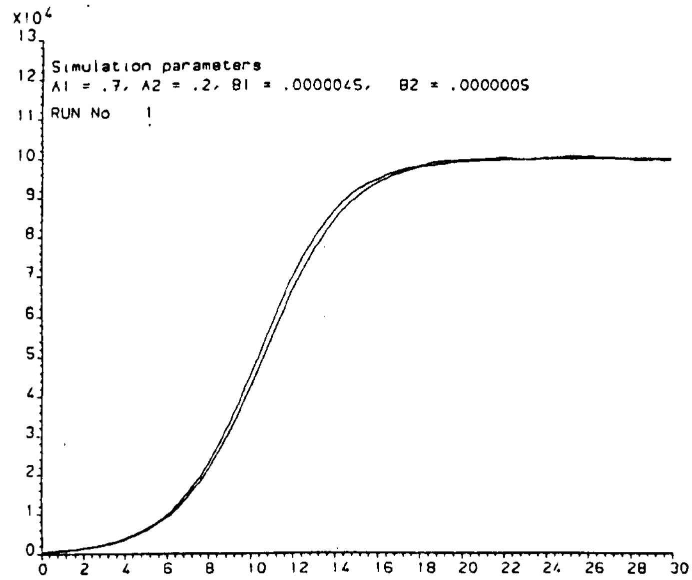
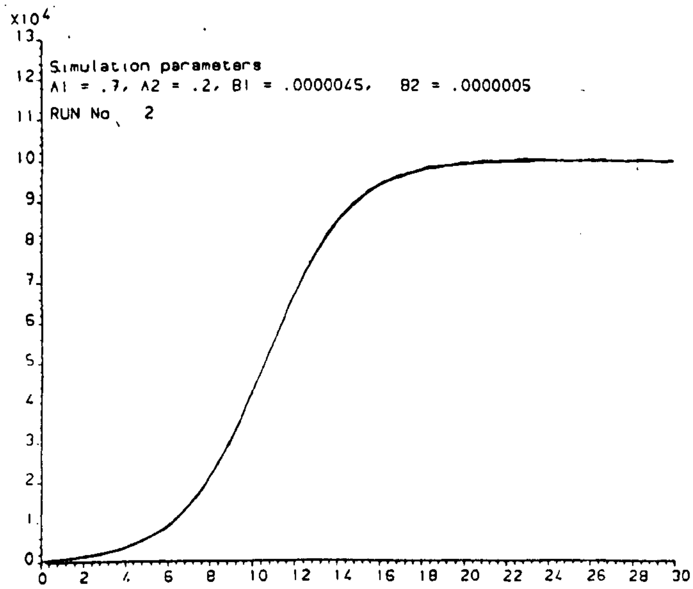
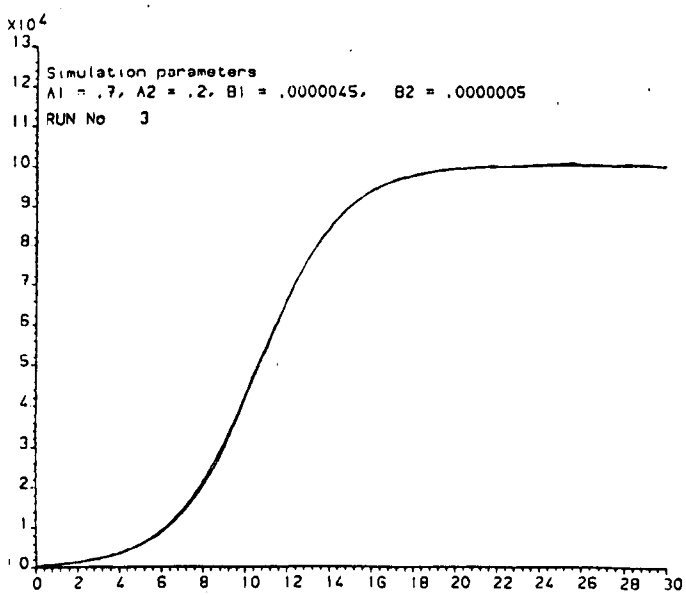


Figure 2.5.1 (continued)







model to these simulations to the underlying parameters upon which the simulation was based. (The results presented are "averaged" findings for 10 simulations for each parameter setting.) The parameter estimates were obtained by Glasbey's routine and by using the NAG 10 Library algorithm E∅4FCF with no correction for correlated residuals.

Reviewing Table 2.5.1 the results can be summarised as:

(i) Glasbey's routine gives slightly closer fits than the NAG routine although all fits are good except for the shortest series.

(ii) Parameter estimates are not significantly different and are good except for shorter series.

Thus it makes little difference which routine is used.

The effect of correcting for correlated residuals is now considered using Glasbey's method and routine. Results are displayed in Table 2.5.2. By considering these results it would seem that the inclusion of an autoregressive error of order one gives the biggest improvement and modelling with a four or five autoregressive component often gives poorer performance. (It is also of note that the coefficients of autoregressive orders higher than one were often not significantly different from zero).

However, the improvement in giving closer fits and more accurate estimates of the parameters is not significant (i.e. the standard errors of the parameter estimates overlapped). Similar results are found when the correction for correlation was applied when actual population data was fitted. Table 2.5.3 displays a summary of fits investigated (more detail on fits of population is given in Section 2.7).

By examining Table 2.5.3 one can see that correcting for

Table 2.5.1: Comparison of Fitting Algorithms

Number of data points	25		20		15		10		5	
	Percentage error									
Exact Parameter Values	Glasbey	NAG	Glasbey	NAG	Glasbey	NAG	Glasbey	NAG	Glasbey	NAG
N_∞ = 100 N_o = 10 C = 0.5	3.6 19.6* 4.5	2.3 13.2 7.8	4.6 20.3* 3.4	3.5 13.7 6.5	+ -2.2 16.2* 11.4	+ -2.8 11.2 13.4	/ -4.4 15.4 13.5	0 -4.8 10.8 15.0	69.0* 3.4 89.5*	68.6* 3.1 85.8*
Variance of Residuals	6.4×10^{-3}	7.4×10^{-3}	5.7×10^{-3}	6.5×10^{-3}	5.85×10^{-3}	6.9×10^{-3}	7.2×10^{-3}	9.5×10^{-3}	3.3×10^{-3}	7.8×10^{-3}
N_∞ = 1000 N_o = 50 C = 0.125	0.83 -1.1 0.8	0.78 -1.5 0.8	0.95 -1.1 0.98	0.92 -1.5 0.80	+ -0.13 -1.5 1.6	+ -0.13 -1.5 1.6	Ω 6.6 -0.2 -0.8	Ω 6.5 -0.2 -1.1	< -5.5 0.9 -1.6	< -5.1 0.9 -1.1
Variance of Residuals	6.7×10^{-4}	7.7×10^{-4}	8.5×10^{-4}	6.1×10^{-4}	7.6×10^{-4}	9.3×10^{-4}	7.4×10^{-4}	1.04×10^{-3}	1.5×10^{-4}	3.7×10^{4}
N_∞ = 10,000 N_o = 200 C = 0.5	+ -1.35 2.0 3.2*	+ -1.36* 1.7 3.2*	+ -1.16* 2.0 3.2	+ -1.17* 1.8 3.2	\$ -1.4 2.0 3.3*	\$ -1.4 1.8 3.2*	≈ -4.2 1.5 3.9*	≈ -4.3 1.34 3.9*	^ -43.1* -0.5 10.8*	^ -42.4* -0.7 10.6
Variance of Residuals	1.9×10^{-4}	2.1×10^{-4}	1.6×10^{-4}	1.8×10^{-4}	2.5×10^{-4}	2.7×10^{-4}	3.1×10^{-4}	3.4×10^{-4}	1.5×10^{-4}	5.9×10^{-4}
N_∞ = 100,000 N_o = 500 C = 0.1	+ 2.0 -7.2* -2.3*	+ 1.9 -7.6* -2.2*	\$ 1.1 -7.4* -2.1	\$ 1.0 -8.0* -2.0	/ 6.7 -6.4* -3.0*	/ 6.3 -6.9* -2.9*	Ω 37.6 -5.2* -4.5*	Ω 33.8 -5.6 -4.2*	> FAILED	> FAILED
Variance of Residuals	5.6×10^{-4}	6.0×10^{-4}	4.8×10^{-4}	4.9×10^{-4}	5.7×10^{-4}	6.7×10^{-4}	6.1×10^{-4}	8.6×10^{-4}	2.4×10^{-4}	2.6×10^{-3}

* Error greater than two standard errors of the parameter estimate

\+ = last observation is 98% of upper asymptote
\$ = last observation is 95% of upper asymptote
/ = last observation is 90% of upper asymptote
Ω = last observation is 83% of upper asymptote
≈ = last observation is 68% of upper asymptote
> = last observation is 32% of upper asymptote
^ = last observation is 28% of upper asymptote
< = last observation is 15% of upper asymptote

Table 2.5.2: The effect of correcting for correlated errors

Description of and parameters on which simulation was based	Percentage error in parameter estimates Order of Error Structure					
	0	1	2	3	4	5
Whole series 25 counts N_∞ = 100 N_0 = 10 c = 0.5 Var.of Residuals	4.6 20.3* 3.3 5.7×10^{-3}	4.3 17.2* 5.8 5.72×10^{-3}	4.5 15.4* 5.6 5.17×10^{-3}	4.5 16.0* 4.8 5.05×10^{-3}	4.0 15.1* 7.7 4.2×10^{-3}	4.1 14.4* 7.4 4.2×10^{-3}
76% of up.asymp. 7 counts N_∞ = 100 N_0 = 10 c = 0.5 Var.of Residuals	-18.8 11.9 27.3 2.26×10^{-3}	14.2 20.2* 15.5 6.69×10^{-3}	-34.0* -10.0* 74.0* 2.64×10^{-3}	-37.8* -13.3* 74.0* 2.26×10^{-3}	Fitting Failed	Fitting Failed
Whole series 66 counts N_∞ = 1000 N_0 = 800 c = 0.125 Var.of Residuals	-1.3 -0.75 8.8* 4.8×10^{-4}	-1.1 -0.5 8.8* 2.4×10^{-4}	-1.1 -0.5 9.1* 2.4×10^{-4}	-1.1 -0.5 8.8* 2.4×10^{-4}	-1.4 -1.5 9.6* 2.2×10^{-4}	Fitting Failed
97.5% of u.a. 69 counts N_∞ = 10,000 N_0 = 500 c = 0.1 Var.of Residuals	0.5 -7.2* -4.7* 2.4×10^{-4}	2.12 -2.4 -7.7* 7×10^{-5}	1.8 -3.3* -7.1* 7×10^{-5}	1.9 -3.2 -7.2* 7×10^{-5}	1.4 -4.2* -6.3* 5×10^{-5}	1.6 -3.7* -6.6* 5×10^{-5}
68% of up.asymp. 40 counts N_∞ = 10,000 N_0 = 500 c = 0.1 Var.of Residuals	7.0* -6.0* -6.6* 3×10^{-4}	13.1 -2.4 -9.7* 1.1×10^{-4}	Fitting Failed	Fitting Failed	Fitting Failed	Fitting Failed
Whole series 159 counts N_∞ = 100,000 N_0 = 5000 c = 0.05 Var.of Residuals	0.6 -0.4* -0.2 3×10^{-5}	0.1 -2.6* -0.5* 2×10^{-5}	0.1 -2.6* -0.5* 2×10^{-5}	0.1 2.9* -0.4 2×10^{-5}	56.2F 68.2F -3.0F 4×10^{-5}	3.8F 24.2F -8.0F 1×10^{-5}
78% of up.asymp. 14 counts N_∞ = 100,000 N_0 = 5000 c = 0.05 Var.of Residuals	5.6* -2.2* -0.1 1.3×10^{-4}	5.5* -2.4* -0.1* 1.2×10^{-4}	5.6* -2.4* -0.1* 1.2×10^{-4}	Fitting Failed	Fitting Failed	Fitting Failed

* = error in excess of two standard errors from the parameter estimates.

Table 2.5.3: The effect of correcting for correlation in fitting+ observed populations

Population and model	Parameters	Estimate of parameters					
		Order of error structure					
		0	1	2	3	4	5
Gt. Britain 1801-1981 19 counts 3 parameter logistic t*-yr-1801	N_∞	62.743 (1.267)	62.497 (1.684)	61.861 (1.423)	61.936 (1.054)	61.770 (1.173)	62.121 (1.071)
	N_0	9.968 (0.433)	9.988 (0.342)	9.901 (0.342)	9.847 (0.237)	9.872 (0.280)	9.861 (0.232)
	c	0.0198 (0.00064)	0.01989(0.00089)	0.02018 (0.00076)	0.02018(0.00055)	0.02019(0.00062)	0.02011(0.00054)
	Var.of residuals	0.456 (0.148)	0.347 (0.113)	0.278 (0.090)	0.216 (0.070)	0.204 (0.066)	0.186 (0.060)
Gt. Britain 1801-1981 19 counts 4 parameter logistic t*=yr-1801	N_∞	56.000 (2.931)	55.980 (3.537)	52.692 (2.405)	51.538 (1.109)	51.416	51.859
	N_0	6.697 (1.317)	6.709 (1.628)	5.494 (1.035)	5.050 (0.455)	5.001	5.146
	c	0.0232 (0.00175)	0.02319(0.0022)	0.02516 (0.00172)	0.02516(0.00082)	0.0261	0.02576
	d	3.927 (1.652)	3.887 (2.013)	5.390 (1.374)	5.978 (0.630)	6.049	5.852
	Var.of residuals	0.383 (0.124)	0.311 (0.101)	0.200 (0.065)	0.0731 (0.0237)	FAIL	FAIL
U.S.A. 1790-1970 19 counts 3 parameter logistic t*-yr-1801	N_∞	309.974(27.178)	417.354(86.703)	299.500 (27.463)	292.094(24.348)	314.076(31.155)	384.441
	N_0	6.093 (0.750)	7.870 (1.779)	5.724 (0.795)	5.531 (0.691)	5.632 (0.668)	9.933
	c	0.02475(0.00133)	0.02158(0.00219)	0.0255 (0.00151)	0.02588(0.00139)	0.02542(0.00139)	0.01854
	Var.of residuals	14.462 (4.692)	8.225 (2.669)	6.424 (2.084)	6.247 (2.027)	6.0126 (1.951)	FAIL
Sweden 1810-1970 42 counts 3 parameter logistic t*=yr-1750	N_∞	20.843 (2.755)	27.188 (10.979)	23.407 (7.382)	22.303 (6.251)	22.015 (5.972)	20.489 (4.532)
	N_0	1.836 (0.034)	1.907 (0.087)	1.882 (0.076)	1.870 (0.070)	1.867 (0.0685)	1.849 (0.059)
	c	0.00926(0.0004)	0.0085 (0.0085)	0.00885 (0.0008)	0.009 (0.0008)	0.009 (0.0008)	0.00924(0.0007)
	Var.of residuals	0.0119 (0.0026)	0.0038 (0.00083)	0.0035 (0.0008)	0.00347(0.0008)	0.0035 (0.0008)	0.0033 (0.0007)

Key: + - unweighted fitting.
FAIL - refers to error structure becoming non-stationary resulting in failure of likelihood algorithm to converge.
Standard errors are in brackets.

correlation has little effect for modelling the population of Great Britain. However, the value of their upper asymptote changes when one autoregressive component is included for the USA and Sweden but returns to a similar level as for the zero order fit when more autoregressive components are included.

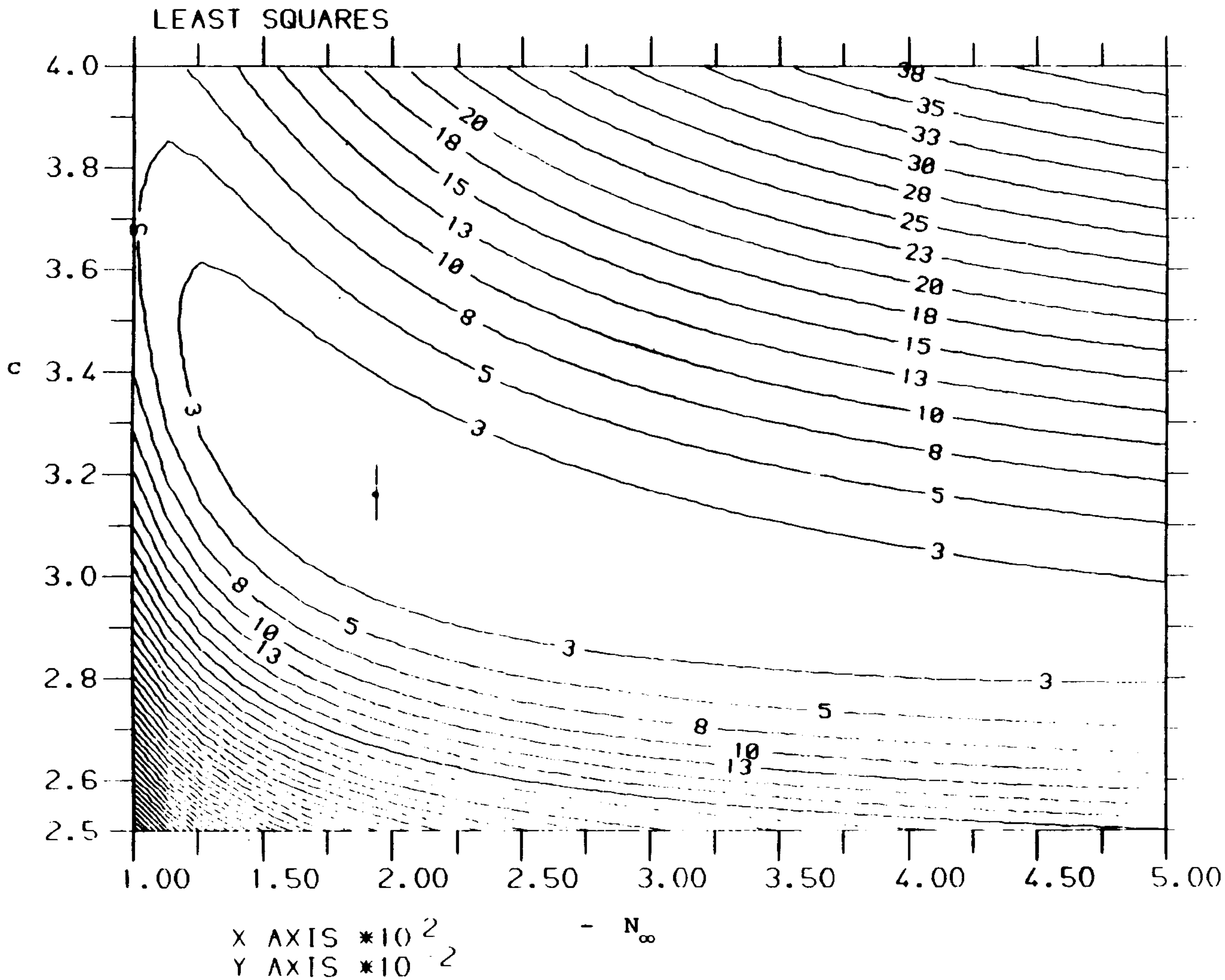
In this initial consideration of fitting sigmoids two anomalies were observed - there is a wave-like pattern to the residuals and there are strong correlations between the parameters, especially between the upper asymptote and the growth parameter which are strongly negatively correlated (Leach, 1981, comments on this correlation). This correlation adds to the problem of estimating parameters as it makes the sum of squares of the residuals flat in the direction of the upper asymptote, as can be observed from Figure 2.5.2.

By introducing an autoregressive error structure to the fitting algorithm little change is made to the wave patterning of the residuals nor to the correlation between parameters.

It would appear that for the type of data that is considered in this thesis little is to be gained from the more complicated procedure of correcting for inherent correlation in the data. Thus to fit the trend models to population data, Glasbey's routine with zero order error structure will be used. This routine requires slight alteration for the case when the data is not spaced at regular intervals. It was found to be simpler to use the NAG10 routine for such cases which gives, as has been shown, similar results.

Various transformations, such as taking roots, reciprocals or natural logarithms were also examined and found to give no significant improvement.

Figure 2.5.2: The % sum of squares surface for the 3-parameter logistic model fit to the USA 1790-1930



2.6 The Fitting Procedure Reconsidered

In the light of the previous three Sections the methodology that will be used to fit models to the population counts is as follows:

(i) Explore the data - as discussed in Section 2.4.

(ii) Fit the model using weighting, i.e. minimise $\sum \left(\frac{P_t - \hat{P}_t}{\hat{P}_t}\right)^2$, and ignore serial correlation (since it is argued in previous Sections that correcting for serial correlation has little effect).

(iii) Assess the quality of the model, by comparing the model to other models, by considering its interpolation and extrapolation abilities, giving value to the demographic appeal of the parameters and closeness of fit.

The models which are compared are:

1. <u>3 Parameter Logistic</u> (3 P.L.)

$$P_t = \frac{N_\infty}{1 + \left(\frac{N_\infty}{N_0} - 1\right)e^{-ct}}$$

2. <u>4 Parameter Logistic</u> (4 P.L.)

$$P_t = 3\text{ P.L.} + d$$

3. <u>3 Parameter Gompertz</u> (3 P.G.)

$$P_t = N_\infty\, e^{be^{-ct}}$$

4. <u>4 Parameter Gompertz</u> (4 P.G.)

$$P_t = 3\text{ P.G.} + d$$

5. <u>Exponential</u> (Geometric)

$$P_t = Ae^{Bt}$$

6. <u>Modified Exponential</u>

$$P_t = A + Be^{ct}$$

7. <u>Linear</u>

$$P_t = A + Bt$$

8. Quadratic

$$P_t = A + Bt + Ct^2$$

9. Cubic

$$P_t = A + Bt + Ct^2 + Ct^3$$

10. Cyclical

$$P_t = A + B \sin(Ct + D).$$

As a consequence of (iii), judgment plays a major role for model selection in this thesis and is favoured rather than the pursuit of greater sophistication and specialisation in modelling.

Before embarking on fitting these models, it is worthwhile to consider if the growth models are of a parametric form that when fitting by least squares methods behaves similarly to a linear model.

The choice of parametric form is important as a model which is close to linear in behaviour is easier to fit, convergence is faster, and one can attach greater weight to the properties of the inverse asymptotic covariance matrix such as the standard errors of the parameter estimates (Ratkowsky, 1983). Models which exhibit strong non-linear behaviour are not amenable to the inference methodology developed for linear models, such as t-tests of parameter estimates.

Oliver (1982) argues that a better parametric form of the logistic is:

$$P_t = \frac{N_\infty}{1 + be^{-ct}}$$

as this reduces correlations between the parameters.

Bates and Watts (1980) developed analytical measures to assess how "non-linear" non-linear models are. This method examines the curvature of the solution surface. There are two measures:

(a) Intrinsic Non-linearity (IN)) which is measured by the curvature of the solution locus around the parameter estimates. If the intrinsic non-linearity is high (greater than the F distribution statistic $(4F_{v_1,v_2,\alpha})^{-1/2}$, then the class of models is inherently non-linear and no reparameterisation will improve the situation.

(b) Parameter Effects Non-linearity (PE) - this is measured by the spacing and parallelism of the parameter lines projected onto the tangent plane to the solution locus. A satisfactory PE is indicated by PE < $(4F_{v_1,v_2,\alpha})^{-1/2}$. PE can be improved by reparameterising.

Ratkowsky (1983) explains these measures further and illustrates the use of this means of assessing non-linearity and comparing different parameterisations in vegetative growth models.

In assessing the contribution to non-linear behaviour, use can be made of Box's percentage bias (a low bias being desirable, see Box (1971), also Ratkowsky (1983)).

In investigating the use of sigmoids to model vegetative growth, Ratkowsky found that the logistic behaved closest to linear behaviour for both IN and PE. However, Ratkowsky did not find the reparameterisation suggested by Oliver to be good as it exhibits significant non-linear parameter effects.

Different parameterisations of the logistic and Gompertz models fitted to the Great Britain population 1801-1981 (19 counts) are presented in Table 2.6.1, along with correlations between parameters, Box's bias on parameters and the non-linearity measures of Bates and Watts. The critical value of $(4F_{3,16,\alpha})^{-1/2}$ is 0.217 at the 1% level and 0.278 at the 5% level. From Table 2.6.1 one can conclude that, although all models have

Table 2.6.1: Re-paramaterisations of Models for the Population of Great Britain

Model	Parameterisation	Parameter Estimates	Standard Errors	Correlations between Parameters	Box's Bias	Intrinsic Non-linearity	Parameter Effects Non-linearity
Logistic	$P_t = \dfrac{N_\infty}{1+\left(\dfrac{N_\infty}{N_o}-1\right)e^{-ct}}$	N_∞ = 62.90	1.353	1.0	0.1046	0.032	0.3547
		N_o = 9.98	0.317	0.68 1.0	0.0002		
		c = 0.0198	0.0007	-0.897 -0.906 1.0	-0.0277		
Logistic (Olivers Model)	$P_t = \dfrac{N_\infty}{1 + be^{-ct}}$	N_∞ = 62.91	1.350	1.0	0.1047	0.032	0.6606
		b = 5.30	0.208	-0.146 1.0	0.1922		
		c = 0.0198	0.0007	-0.896 0.530 1.0	0.0407		
Logistic	$P_t = \dfrac{N_\infty}{1+e^{(b-ct)}}$	N_∞ = 62.93	1.344	1.0	0.1050	0.032	0.6622
		b = 1.67	0.033	-0.1521 1.0	0.0897		
		c = 0.0198	0.0007	-0.896 0.536 1.0	0.0407		
Logistic	$P_t = \dfrac{c/a}{1+e^{-c(t-b)}}$	a = 0.0032	0.00002	1.0	0.0529	0.032	0.7478
		b = 83.80	2.75	-0.929 1.0	0.1462		
		c = 0.0199	0.0007	0.985 -0.865 1.0	0.0409		
Logistic	$P_t = \dfrac{1}{a+be^{-ct}}$	a = 0.0159	0.0004	1.0	-0.0563	0.032	0.3425
		b = 0.084	0.0043	0.6372 1.0	0.1317		
		c = 0.0198	0.0007	0.8967 0.8819 1.0	0.0407		
Logistic	$P_t = \dfrac{N_\infty}{1 + bc^t}$	N_∞ = 62.90	1.350	1.0	0.1046	0.032	0.6551
		b = 5.30	0.208	-0.1452 1.0	0.1922		
		c = 0.980	0.0007	0.8962 -0.5294 1.0	-0.0008		
Logistic	$P_t = \dfrac{1}{a + bc^t}$	a = 0.0159	0.00034	1.0	-0.0563	0.032	0.3381
		b = 0.0843	0.0043	0.6361 1.0	0.1317		
		c = 0.9804	0.0007	-0.8965 -0.8814 1.0	-0.0008		
Logistic	$P_t = \dfrac{1}{a + b^t e^c}$	a = 0.0159	0.00034	1.0	-0.0563	0.032	0.3738
		b = 0.980	0.0007	-0.8971 1.0	-0.0008		
		c = -2.47	0.04145	0.6383 -0.8822 1.0	-0.0278		
Gompertz	$P_t = N_\infty e^{be^{-ct}}$	N_∞ = 78.95	4.29	1.0	0.5346	0.0356	2.4955
		b = -2.14	0.044	-0.4278 1.0	0.2460		
		c = 0.010	0.0007	-0.9677 0.251 1.0	0.0756		
Gompertz	$P_t = N_\infty e^{-e^{b-ct}}$	N_∞ = 77.93	4.34	1.0	0.5435	0.0356	2.5458
		b = 0.757	0.021	0.4387 1.0	0.2998		
		c = 0.0102	0.00075	-0.9681 -0.2283 1.0	0.0756		
Gompertz	$P_t = N_\infty e^{bc^t}$	N_∞ = 78.45	4.33	1.0	0.5579	0.0358	2.5884
		b = -2.14	0.044	-0.4349 1.0	0.2519		
		c = 0.99	0.0007	0.9680 -0.2236 1.0	-0.0008		
Gompertz	$P_t = e^{a+bc^t}$	a = 4.36	0.056	1.0	0.0871	0.0358	1.6565
		b = -2.13	0.044	-0.4315 1.0	0.2479		
		c = 0.99	0.0007	0.9677 -0.2193 1.0	-0.0008		
Logistic	$P_t = \dfrac{N_\infty}{1+\left(\dfrac{N_\infty}{N_o}-1\right)e^{-ct}} + d$	N_∞ = 56.25	3.43	1.0	0.6137	0.0572	3.5073
		N_o = 6.77	1.54	0.9709 1.0	2.7016		
		c = 0.023	0.002	-0.9836 -0.9740 1.0	-0.4772		
		d = 3.85	1.92	0.9509 -0.9842 0.9304 1.0	-4.8422		
Logistic	$P_t = \dfrac{1}{a + bc^t} + d$	a = 0.0177	0.0011	1.0	-0.2631	0.0571	10.4238
		b = 0.129	0.032	0.9677 1.0	2.5505		
		c = 0.977	0.002	-0.9829 -0.9716 1.0	-0.0003		
		d = 3.79	1.88	0.9490 0.9837 -0.9277 1.0	-4.9658		

t = year - 1801

no adverse intrinsic non-linearities, there are no satisfactory parameterisations. However, the best forms are logistics. The form $P_t = \dfrac{N_\infty}{1 + \left(\dfrac{N_\infty}{N_0} - 1\right)e^{-ct}}$ is preferred to $P_t = \dfrac{N_\infty}{1 + bc^t}$ as its parameters are directly interpretable in demographic terms.

Another observation is that there does not seem to be much relationship between correlations between parameters and satisfactory linear behaviour, and Oliver's (1982) parameterisation is not the most "linear" of the logistic parameterisations.

Gompertz models are far from linearity when the parameter effects linearity is assessed. The four-parameter models also exhibit more severe non-linear behaviour - this may arise due to the interaction between the parameters N_0 and d.

Consideration is also made of the linear behaviour of models fitted to the Swedish and USA populations. The models fitted to USA and Sweden up to 1980 fail the linearity tests, possibly due to increased variability of the residuals from the fit; this point is reconsidered in the next Section.

The models fitted to the USA 1790-1930 both exhibit satisfactory linear behaviour at the 1% level.

Therefore, in this thesis the model of parametric form:

$$P_t = \frac{N_\infty}{1 + \left(\dfrac{N_\infty}{N_0} - 1\right)e^{-ct}}$$

will be used where possible. The four-parameter version of this, i.e. with an offset d, will also be considered, but the cost of greater non-linear behaviour must be recognised. This cost will result in greater difficulties in fitting and judging model and parameter estimate viability.

Analysis of Tables 2.6.1 and 2.6.2 supports points made by Ratkowsky (1983) that:

(i) it is false to think that high correlations between parameters is either diagnostic or indicative of non-linear behaviour;

(ii) whether or not a parameter appears linearly in a non-linear regression model bears no relationship to its behaviour in estimation;

(iii) parameter effects are greater than intrinsic non-linearities.

Finally, to comment on the methodology of Bates and Watts (1980): the investigation carried out in this Section drew attention to the approach being computationally cumbersome as it requires calculation of second order partial derivatives, since the method was found to be very susceptible to errors such as incorrect signs in the partial derivatives. The comments made in the discussion of Bates and Watts' (1980) paper are echoed in that, generally to use these measures of non-linearity it is highly desirable that one incorporates the use of a computer algebraic package to produce the code for the derivatives.

Table 2.6.2: Re-paramaterisations of Models for different Populations

Model	Parameterisation	Parameter Estimates	Standard Errors	Correlations between Parameters	Box's Bias	Intrinsic Non-linearity	Parameter Effects Non-linearity
Sweden 1810-1980	$P_t = \frac{N_\infty}{1+\left(\frac{N_\infty}{N_0}-1\right)e^{-ct}}$	N_∞ = 15.91 N_0 = 1.48 c = 0.0105	1.377 0.05 0.0005	1.0 0.862 1.0 -0.9709 -0.9536 1.0	1.2521 0.0187 -0.4144	0.0250*	1.0167
Sweden 1810-1980	$P_t = \frac{1}{a + bc^t}$	a = 0.063 b = 0.644 c = 0.99	0.006 0.0217 0.0005	1.0 0.8119 1.0 -0.9752 -0.9160 1.0	-0.4171 0.1752 -0.0003	0.0250*	1.1820
U.S.A. 1790-1980	$P_t = \frac{N_\infty}{1+\left(\frac{N_\infty}{N_0}-1\right)e^{-ct}}$	N_∞ = 354.10 N_0 = 6.66 c = 0.0235	32.98 0.80 0.0013	1.0 0.7897 1.0 -0.9154 -0.9632 1.0	1.1421 0.3541 -0.3064	0.0547*	0.6907
U.S.A. 1790-1980	$P_t = \frac{1}{a + bc^t}$	a = 0.00282 b = 0.147 c = 0.98	0.0003 0.0177 0.0013	1.0 0.7853 1.0 0.9158 0.9611 1.0	-0.3109 1.2258 -0.0033	0.0546*	1.3001
U.S.A. 1790-1930	$P_t = \frac{N_\infty}{1+\left(\frac{N_\infty}{N_0}-1\right)e^{-ct}}$	N_∞ = 198.88 N_0 = 4.03 c = 0.031	5.33 0.14 0.0005	1.0 0.7958 1.0 -0.9174 -0.9645 1.0	0.0980 0.0269 -0.0264	0.0156*	0.2038*
U.S.A. 1790-1930	$P_t = \frac{1}{a + bc^t}$	a = 0.00503 b = 0.243 c = 0.97	0.00013 0.009 0.0005	1.0 0.7404 1.0 -0.9174 -0.9623 1.0	-0.0267 0.0973 -0.0004	0.0156*	0.3537

$* < \frac{1}{2\sqrt{F_{v,p,5\%}}}$

2.7 Some Populations - Modelled

In this Section the results of fitting the models as described in previous Sections are presented. Examples of populations and sub-populations from a variety of countries are considered and the results presented in separate sections for populations rather arbitrarily classified into developed, emerging, underdeveloped and island nations. More detail is given on populations which have been examined in greater depth - these are Great Britain, USA, France, Sweden and Japan from the developed nations, Brazil, Mexico and India from the emerging nations, and Bangladesh and Kenya from the developing nations.

For all populations exploratory analysis has been carried out and comparisons made of different trend models. Specific consideration is made of modelling the sub-populations of Great Britain and also Quebec which is of interest as it represents a population whose recent history has involved very little migration. The maximum time period of these models is limited by the data available and is often less than 200 years. However, the estimates of the English population made by Wrigley and Scofield (1981) offer a longer time series which is examined.

In this Section no mention will be made of "awkward" populations such as Ireland which has declined, nor the Netherlands which seems to have undergone successive short phases of growth and stagnation. These populations are investigated in Section 2.8.

The data to be modelled has, where possible, been taken from National Yearbook records of census counts. These are fairly reliable for developed nations but less so for developing nations. When the data was not available or was thought to be of dubious quality, reference was made to the "United Nations Demographic Yearbook" for United Nations estimates.

2.7.1 Developed Nations

2.7.1a Great Britain 1801-1981

For Great Britain the population has been enumerated regularly in censuses every ten years since 1801 (apart from during the Second World War in 1941 - for this data an estimate of population size taken from OPCS yearly estimates is used).

Exploratory Examination

1. Spline Analysis - the results of spline smoothing are displayed in Figure 2.7.1.1. Figure 2.7.1.1 illustrates a spline fitted to the data (with an optimal smoothing parameter of 47.9m) and confidence intervals. One can observe that the spline intersects all the points and the confidence interval is evenly narrow which indicates that there are no outliers in the data and variance is constant.

Figure 2.7.1.1b displays the derivative of the spline. This does not conform to the idealised "∩" shape as there is another lower peak shortly after the Second World War (the 'baby boom'). This may suggest escalation is required from around 1940. But there is uncertainty as the renewed faster growth is short lived and the previous path of declining growth rates is resumed by 1981.

If this short spurt of growth is neglected, the peak rate of growth, which corresponds to the point of inflection, is encountered just after midway along the data, around 1900.

This (from Figure 2.7.1.1a) suggests an upper asymptote for the logistic model of approximately 74m (since $N_\infty = 2 \times$ population at point of inflection). Also the initial growth rate, γ, (from

Figure 2.7.1.1: Spline Analysis Great Britain

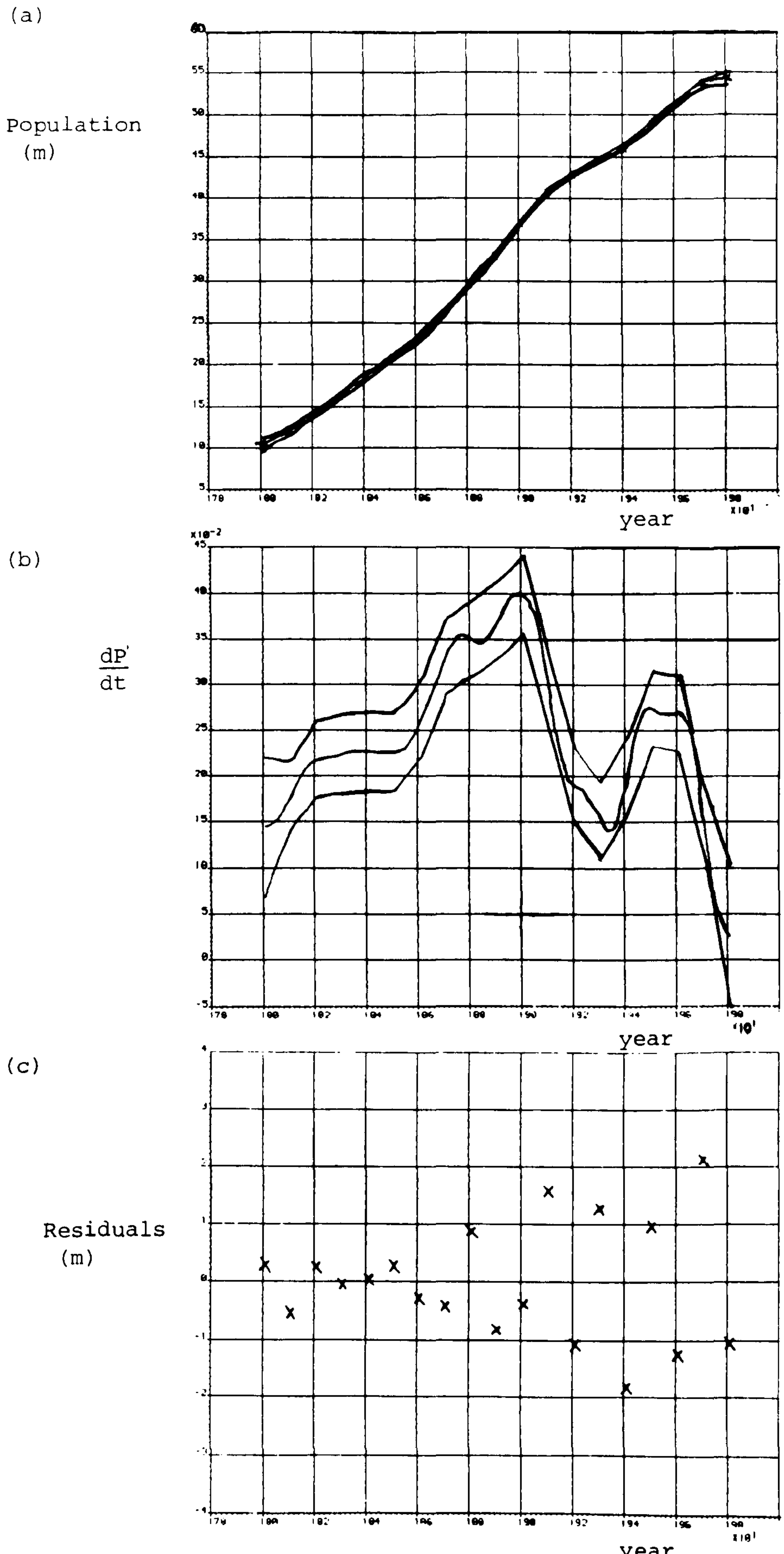


Figure 2.7.1.1b) is between 0.007 and 0.022 (since $\gamma = dN_0/dt_0/N_0$), and may be approximated as 0.0145.

This gives the model:

$$N_t = \frac{74}{1 + (\frac{74}{10} - 1)e^{-0.0145t*}}$$

where t* + 1801 = 1801, 1811, 1821, ..., 1977, 1981.

Since the symmetry is well formed and growth is slowing, the upper asymptote may be encountered in the near future.

Figure 2.7.1.1c illustrates that the variance of the standardised residuals show a tendency to increase, which suggests the need for weighting or transformations when the data is fitted.

2. Cusum - is close to the ideal inverted 'U' and suggests that sigmoids should be used with no escalation (see Figure 2.7.1.2).

3. Time Series Representation - there is insufficient data for Box-Jenkins modelling.

The model $\frac{1}{N_t} = (1 + b)\frac{1}{N_{t-1}} - b\frac{1}{N_{t-2}} + \epsilon_t$ is fitted. From MINITAB the model

$$\frac{1}{N_t} = \underset{(0.1754)}{1.72} \frac{1}{N_{t-1}} - \underset{(0.15465)}{0.734} \frac{1}{N_{t-2}}$$ fitted well,

(0.1754) (0.15465) (standard errors are in brackets)

which suggests that b is around 0.72 and a logistic is an appropriate model. Figure 2.7.1.3a displays the residuals from this model and, from observing this figure, one can speculate that the initial 1801 figure is rather high to be modelled by a logistic model and should be omitted. Doing this does produce a more acceptable residual plot (Figure 2.7.1.3b), but changes the estimate of b to 0.68.

4. Window Analysis - using logistics as models, no evidence to expect escalation or de-escalation could be detected (see Figure 2.7.1.4). (Leach (1981) considered the three parameter logistic and the forward moving window, and found no suspicion of

Figure 2.7.1.2: Cusum for Great Britain

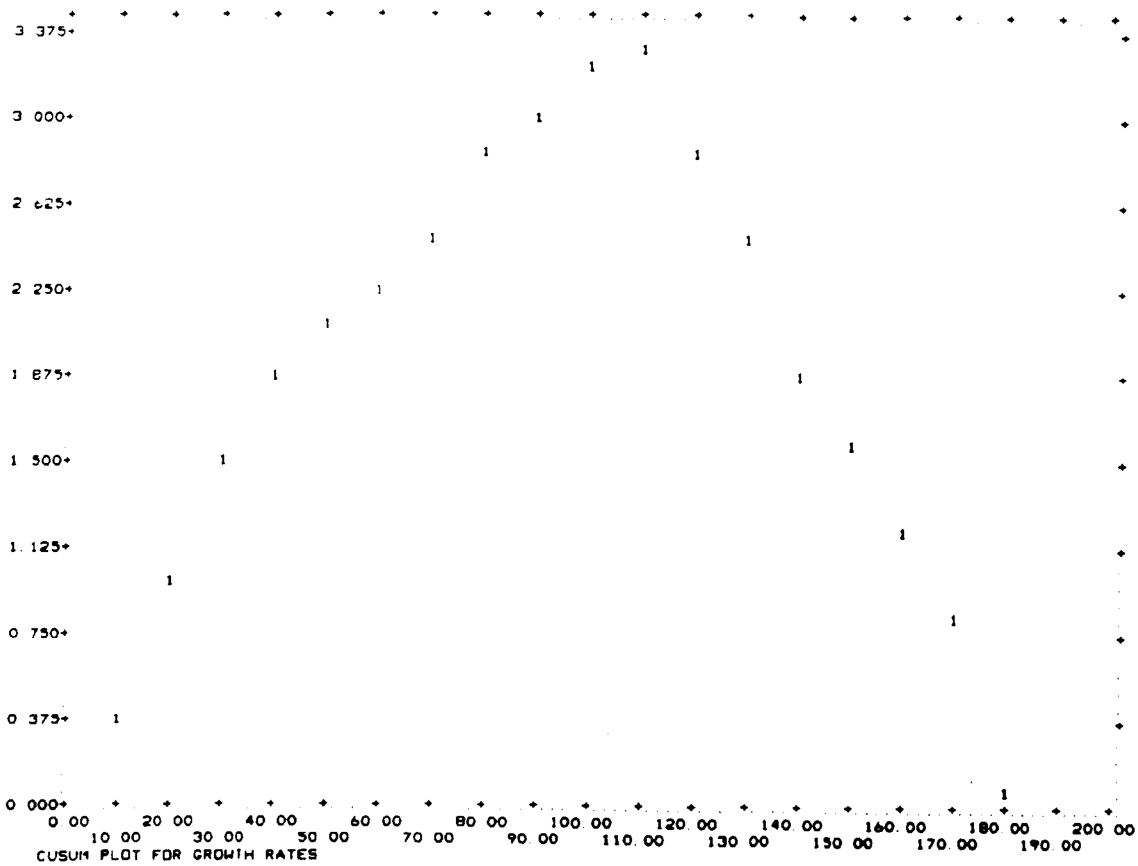
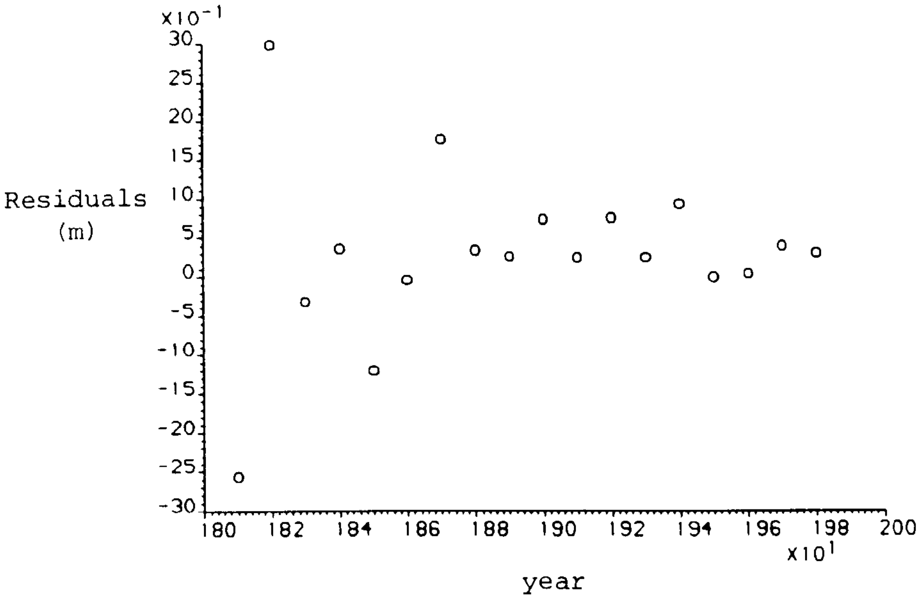


Figure 2.7.1.3: Residuals from 'Linear' model

(a) 1801-1981



(b) 1811-1981

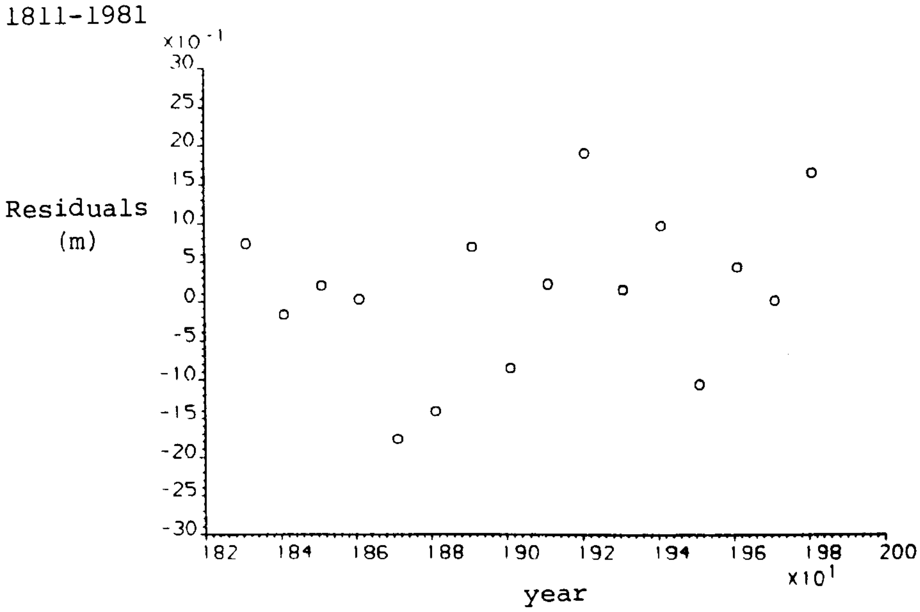
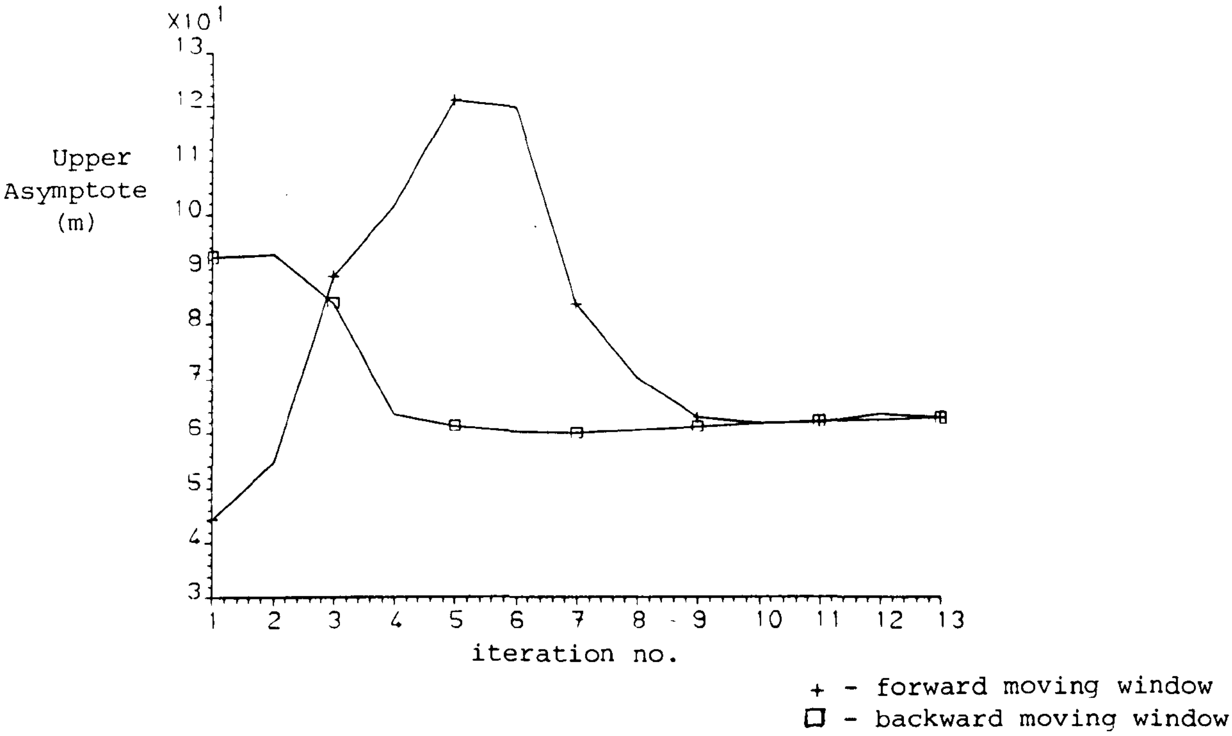


Figure 2.7.1.4a: Window analysis - Great Britain - 3-parameter logistic



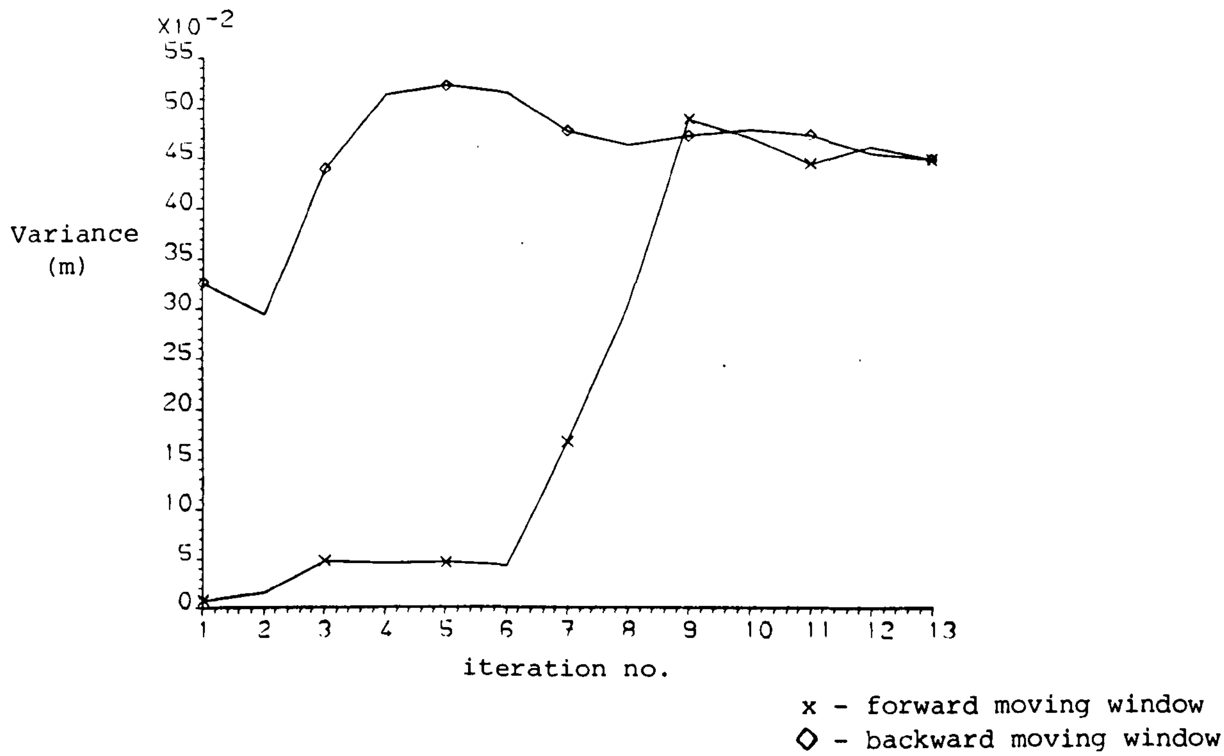
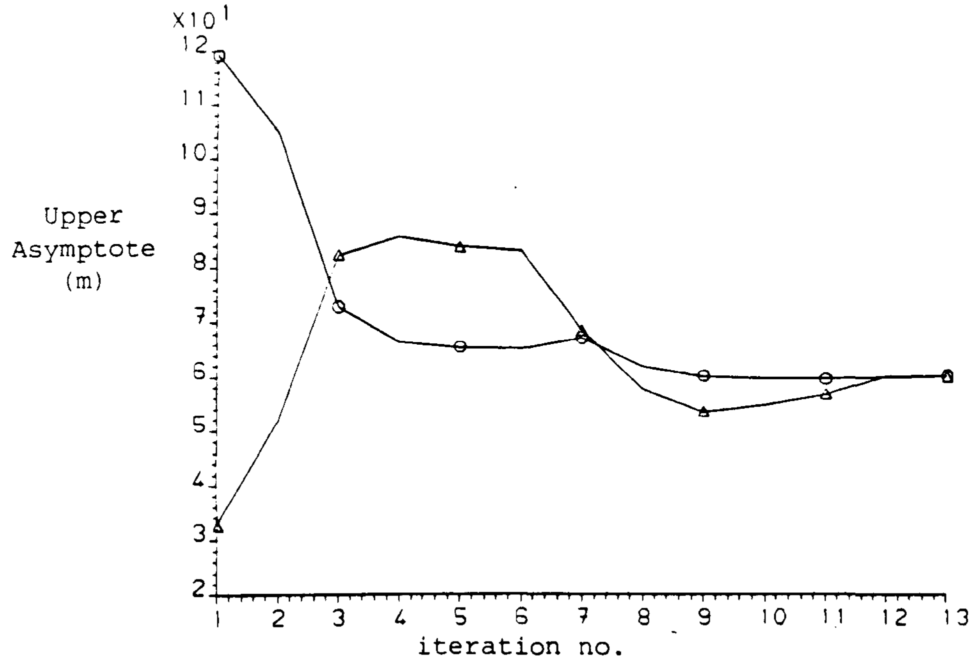
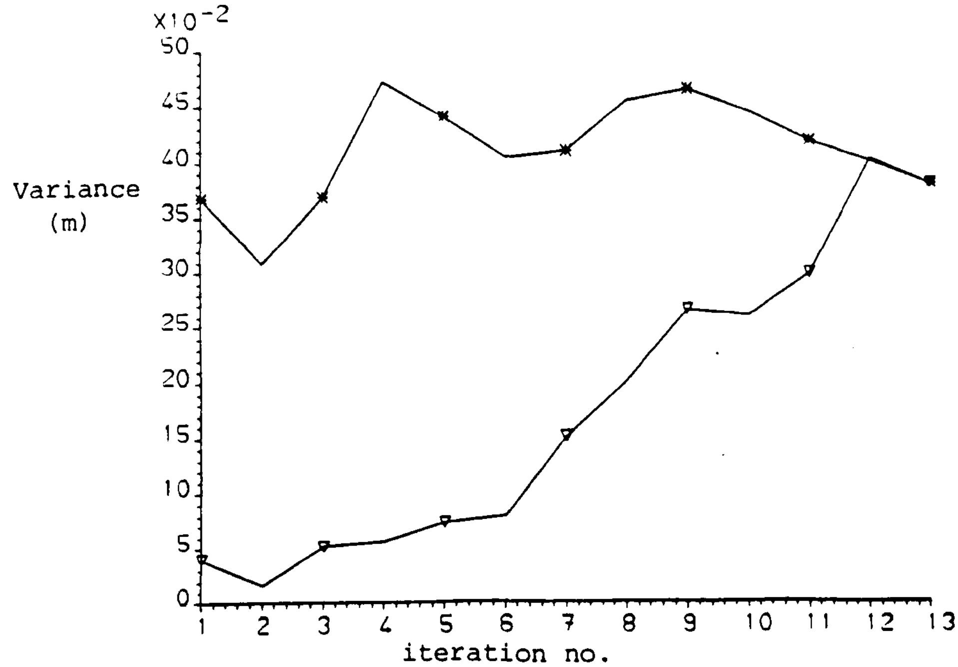


Figure 2.7.1.4b: Window analysis - Great Britain - 4-parameter logistic



△ - forward moving window
o - backward moving window



∇ - forward moving window
* - backward moving window

escalation.)

Models Compared

The simple trend models discussed in Section 2.2 are compared to determine the "best" model, "best" being determined by:

(a) the model that describes the data most closely;

(b) lowest number of parameters;

(c) ability to predict over a period of fifty years. (In this analysis the models are fitted to the period 1801-1931 and the models extrapolated over the period 1941 to 1981 and compared with observations).

Fitting over the period 1801 to 1931 (14 data points), using the computer program PTRENDS (with weighting) gives the models described in Table 2.7.1.1. From this table the logistics appear comparatively good models in terms of lowest likelihood and sum of squares of residuals (SSR). They outweigh the benefit of, say, the geometric model with its fewer parameters even if a cost is made for additional parameters. (A simple cost function to take into account models with different numbers of parameters would be to multiply SSR by the number of parameters minus one.) So the SSR for the four parameter logistic becomes 11.442m and that of the ordinary logistic increases to 9.35m. The costing is shown in the last column of Table 2.7.1.1 in indexed form, the lowest cost having unity as an index.

The logistics still appear as the "best" models, and the four parameter logistic does not seem to be significantly better than the three parameter logistic. (Also note that the fourth parameter of the four parameter logistic is not statistically significant.) An upper asymptote of around 76.7 ± 10m is suggested, which concurs with the spline estimate of the asymptote.

Table 2.7.1.1: Models of the Population of Great Britain 1801-1931

(* = year - 1801)

Model	Parameter	Standard errors	Variance of residuals	Sum of squares of residuals (RSS)	Indexed costed RSS
3 Parameter Logistic	N_∞ = 76.67 N_0 = 10.44 c = 0.0174	4.92 0.12 0.0006	4.7×10^{-4} (10^{-4})	4.675	1.0
4 Parameter Logistic	N_∞ = 66.32 N_0 = 8.51 c = 0.020 d = 2.02	9.45 1.76 0.003 1.85	2.5×10^{-4} (10^{-4})	3.814	1.22
3 Parameter Gompertz	N_∞ = 169.97 b = -2.81 c = 0.006	Failed to converge	4.3×10^{-4}	-	-
4 Parameter Gompertz	N_∞ = 93.84 b = 2.896 c = 0.0096 d = 5.34	24.38 0.126 0.0021 1.83	2.9×10^{-4} (10^{-4})	4.890	1.57
Modified Exponential	A = -18.40 B = 28.68 C = 0.0064	4.62 4.50 0.0007	5×10^{-4} (2×10^{-4})	10.430	2.23
Geometric	A = 11.30 B = 0.0115	0.28 0.0003	2.4×10^{-3} (1.27×10^{-3})	44.722	4.78
Linear	A = 9.52 B = 0.256	0.40 0.009	3.37×10^{-3} (1.27×10^{-3})	34.690	3.71
Quadratic	A = 10.38 B = 0.168 C = 0.00087	0.18 0.009 0.00009	4.1×10^{-4} (1.5×10^{-4})	8.737	1.87
Cubic	A = 10.53 B = 0.137 C = 0.0017 D = -0.000005	0.17 0.017 0.0004 $< 10^{-5}$	3.1×10^{-4} (1.2×10^{-4})	5.117	1.64

Table 2.7.1.2: Percentage error of Forecasts made in 1931 for Great Britain

Year	Observed	3 Parameter Logistic	4 Parameter Logistic	3 Parameter Gompertz	4 Parameter Gompertz	Modified Exponential	Geometric	Linear	Quadratic	Cubic
1941	46.000	6.970	5.837	8.132	7.179	11.853	22.874	-1.332	10.872	7.264
1951	48.854	6.819	5.043	9.237	7.367	14.704	29.797	-1.851	13.012	7.535
1961	51.284	7.214	4.728	11.198	8.243	18.797	38.716	-1.505	16.206	8.479
1971	53.979	6.670	3.464	12.464	8.306	22.512	47.852	-1.676	18.850	8.568
1981	54.286	10.450	6.363	18.619	12.910	32.040	64.934	2.488	26.896	13.144
Absolute Departures										
Sum of squared departures		80.211	34.587	214.770	110.948	624.439	2626.587	4.432	451.210	115.927
M.A.D.		4.005	2.630	6.554	4.711	11.175	22.920	0.941	9.500	4.815

Table 2.7.1.2 compares the prediction ability of these models. Examination of this reveals that the linear model gave by far the most accurate predictions over the next fifty years. This must be assumed as a chance event since the linear model performed badly when fitted to data. Of the other models the logistics gave the closest predictions - but overestimated, the four parameter logistic performing better than the three parameter form.

The models were then refitted to the period 1801 to 1981 (19 data points). A summary of these fits is given in Table 2.7.1.3. These results also show the logistic models to be the "best" descriptors of the Great British population over the period 1801 to 1981. An upper asymptote of approximately 65 ± 3.13m is given by the three parameter logistic model and 60.7 ± 5.0m is suggested by the four parameter logistic model. Since the four parameter logistic model has a better track record for prediction (see Table 2.7.1.2) and allows for previous cycles of growth (the lower asymptote being 3.3 ± 1.0m), the four parameter logistic is chosen as the preferred model. It is noticeable that the asymptote for the fit up to 1981 is significantly lower than the fit up to 1931, for whatever logistic is chosen. This speculatively suggests that the effect of renewed growth around 1940 to late 1950s (see Figure 2.7.1.1b) does not have much effect on the ultimate population of Great Britain. In terms of window analysis (Leach, 1981), this renewed growth can be thought of as failing to counter the falling trend of the upper asymptote.

Figure 2.7.1.5 compares the estimates given by the four parameter logistic and the cubic spline with its 95% confidence limits. From Figure 2.7.1.5 one can see that the four parameter

Table 2.7.1.3: Models of the Population of Great Britain 1801-1981

(* = year - 1801)

Model	Parameter	Standard errors	Variance of residuals	Sum of squares of residuals (RSS)	Indexed Costed RSS
3 Parameter Logistic	N_∞ = 65.06 N_0 = 10.30 c = 0.0188	1.56 0.14 0.0005	4.0×10^{-4} (1.3×10^{-4})	10.012	1.0
4 Parameter Logistic	N_∞ = 57.38 N_0 = 7.24 c = 0.0223 d = 3.32	2.50 0.91 0.0014 1.01	2.9×10^{-4} (1×10^{-4})	7.369	1.10
3 Parameter Gompertz	N_∞ = 93.09 b = -2.224 c = 0.0083	7.42 0.07 0.0006	9.2×10^{-4} (3×10^{-4})	23.034	2.30
4 Parameter Gompertz	N_∞ = 61.22 b = -3.03 c = 0.01376 d = 7.65	3.69 0.18 0.0011 0.80	4.0×10^{-4} (1.3×10^{-4})	8.71	1.31
Modified Exponential	A = -54.51 B = 64.70 C = 0.00032	Failed to converge	2.19×10^{-3}	-	-
Geometric	A = 12.63 B = 0.0093	0.60 0.0005	1.16×10^{-2} (3.76×10^{-3})	386.897	19.33
Linear	A = 9.50 B = 0.257	0.33 0.006	2.57×10^{-3} (8.3×10^{-4})	37.804	1.89
Quadratic	A = 9.94 B = 0.22 C = 0.003	0.34 0.015 0.0001	1.84×10^{-3} (6×10^{-4})	52.402	5.23
Cubic	A = 10.59 B = 0.122 C = 0.0022 D = -0.00001	0.18 0.0133 0.00022 0.000001	3.7×10^{-4} (1.2×10^{-4})	8.779	1.32

logistic model comes close to the benchmark of the spline, only falling significantly outside the 95% confidence limits on two occasions (1911 and 1941 (the estimated count)).

The conclusion then is that a model of the form:

$$N_t = \frac{57}{1 + \left[\frac{57}{7.2} - 1\right] e^{-0.022t*}} + 3 \qquad t* = \text{year} - 1750$$

gives a good description of population size in Great Britain between 1801 and 1981. (Omitting the 1801 count as suggested by the exploratory investigation of the linear model only marginally improved the fit and the parameter estimates were similar, so the whole series was modelled.)

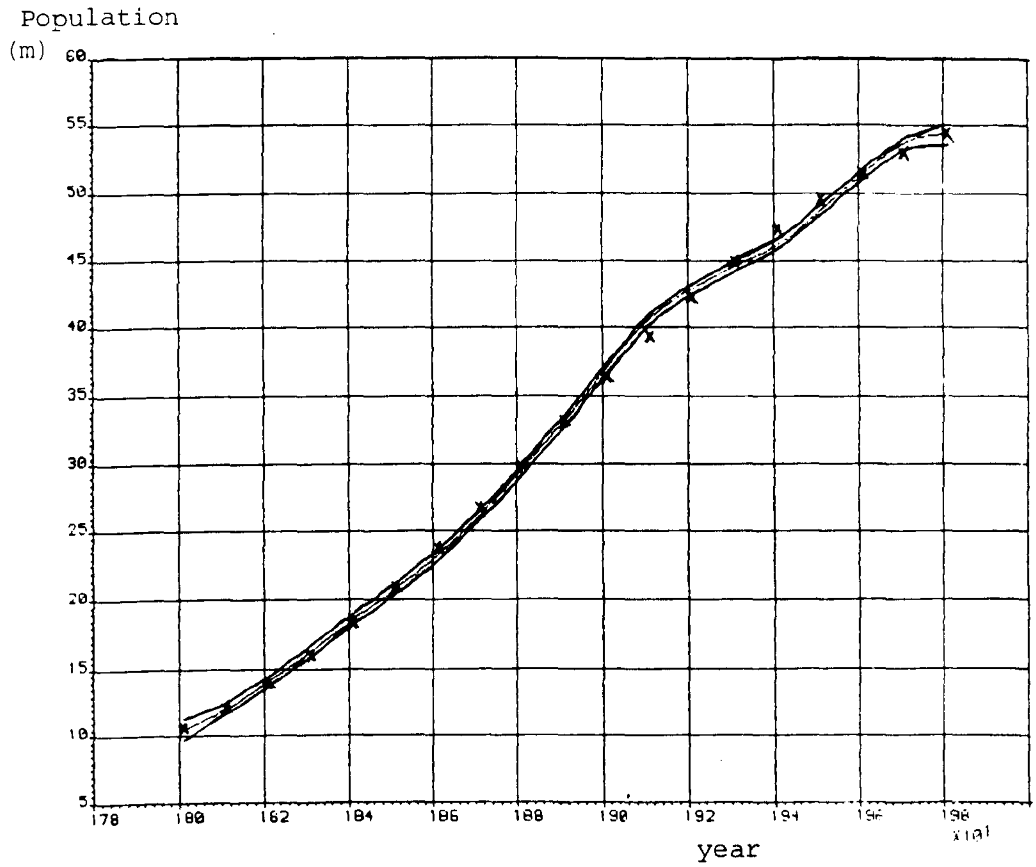
2.7.1b Sweden 1750-1980

This time series is one of the longest available containing forty-five points. After 1810 Swedish censuses have been taken regularly every five years.

Exploratory examination of this data set indicates similarities with the growth rate of the British population; this spline analysis suggests an upper asymptote of around 13m. This, it is noted, does not imply similar causes and consequences for population change; these models are population specific as will be illustrated in Chapter 5.

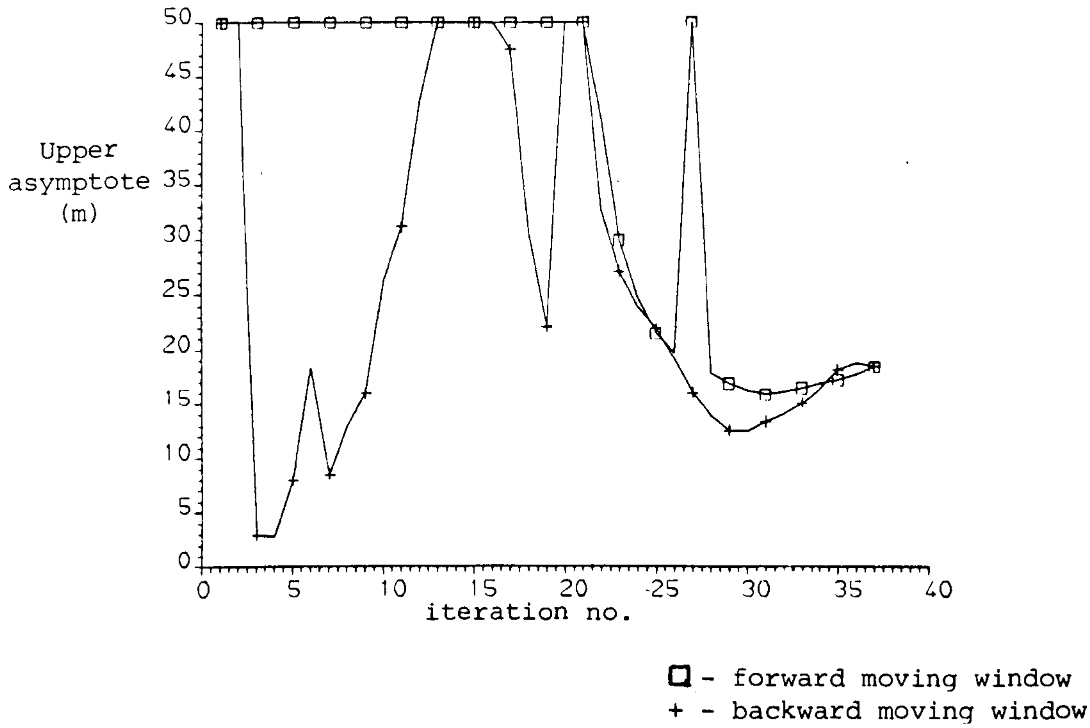
The results of the window analysis based on the three and four parameter logistic models are shown in Figure 2.7.1.6. The forward enlarging window shows slow growth of the upper asymptote and a sharp rise in the sum of squares of residuals after 1930. The backwards enlarging window indicates a rapid rise of the variance for the four data points prior to 1810. Thus, in modelling Sweden's population points prior to 1810, which may be of questionable reliability, will not be included.

Figure 2.7.1.5: Model for Great Britain compared to the spline



x = fitted point

Figure 2.7.1.6a: Window analysis for Sweden - 3-parameter logistic



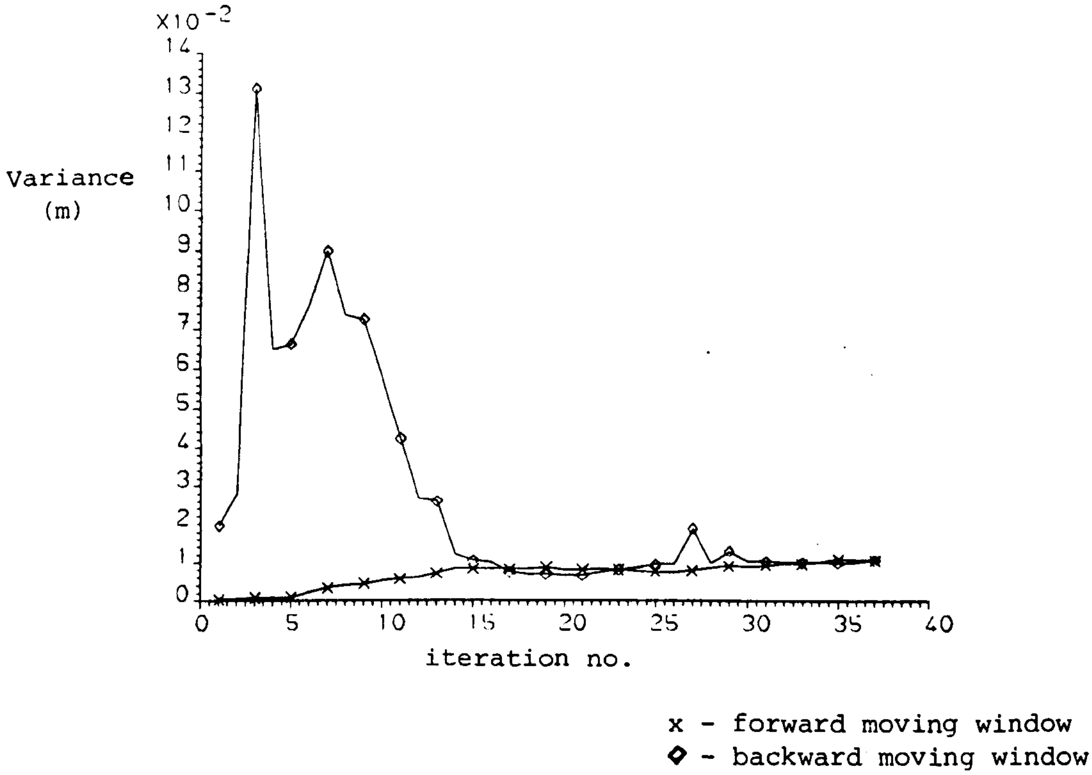
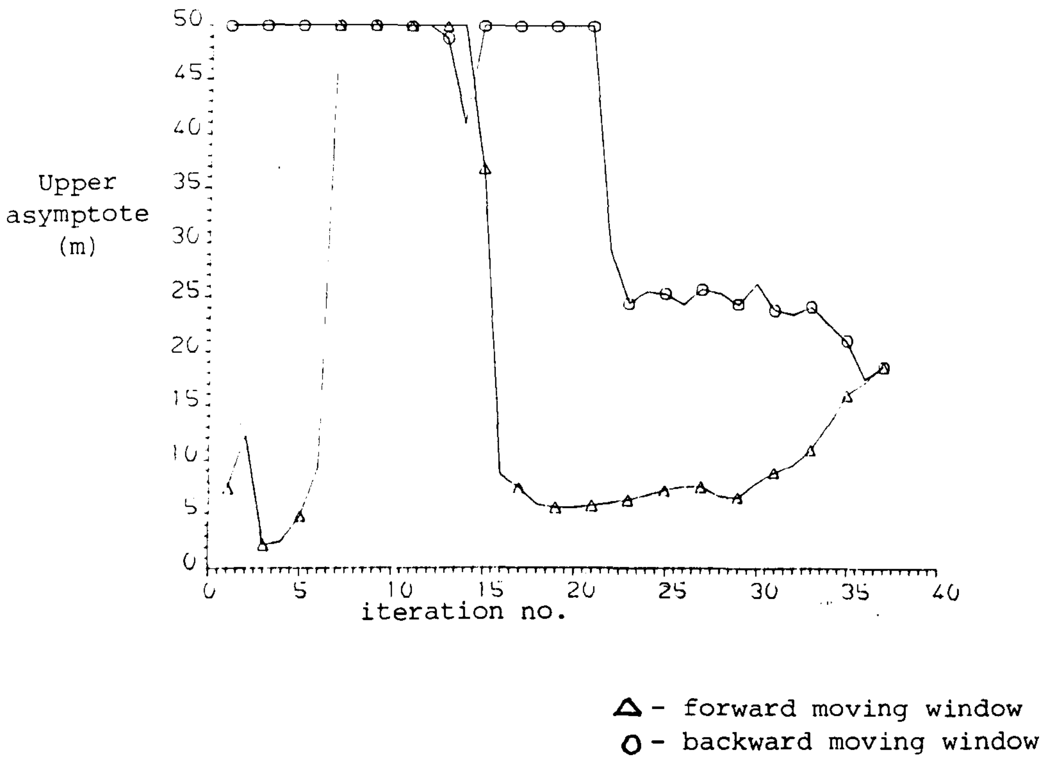
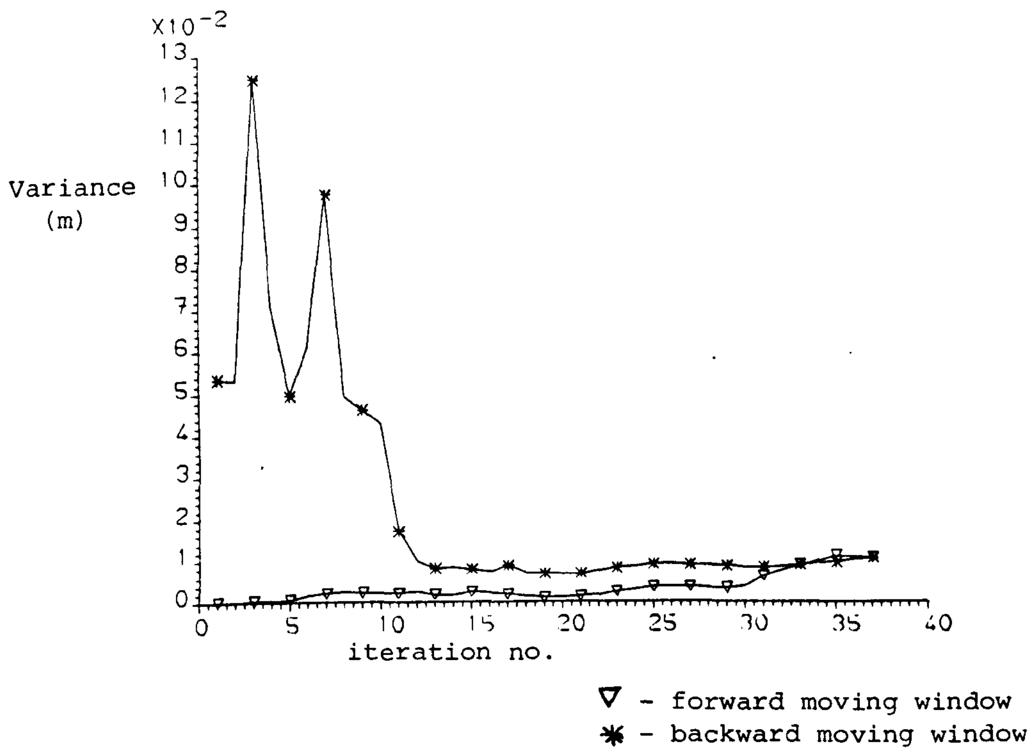


Figure 2.7.1.6b: Window analysis for Sweden - 4-parameter logistic





When the trend models are fitted to the census counts over the period 1810 to 1930 all sigmoids fit well. The three parameter logistic is the best model with parameters:

$$N_\infty = 9.727\ (0.412)$$

$$N_0 = 1.17\ (0.03) \quad t^* = \text{year} - 1750$$

$$\alpha = 0.014\ (0.0005)$$

$$\text{Variance} = 1.3 \times 10^{-4}$$

The predicting performance of this model is good up to 1980 (less than 10% error) but it is marginally bettered by the linear and three parameter Gompertz models. The quadratic performs particularly well (see Table 2.7.1.4).

Fitting the whole period 1810 to 1981 shows that the sigmoids are no longer as good fits as they were for the shorter period. The quadratic model appears to be a better model. This finding, together with results from the exploratory analysis suggests that there may be a need to account for escalation after 1930. Examination of the derivative of the spline (see Figure 2.7.1.7) gives possible escalation to have occurred between 1930 and 1950.

Yearly estimates of the Swedish population for the period 1950 to 1981 were examined to give more data so that a further assessment of the logistic models after escalation can be made. The models obtained are:

Three Parameter Logistic

Parameters	Standard Errors	Sum of Squares of Residuals
$N_\infty = 9.66$	0.34	
$N_0 = 0.067$	0.051	0.0008
$c = 0.0297$	0.004	

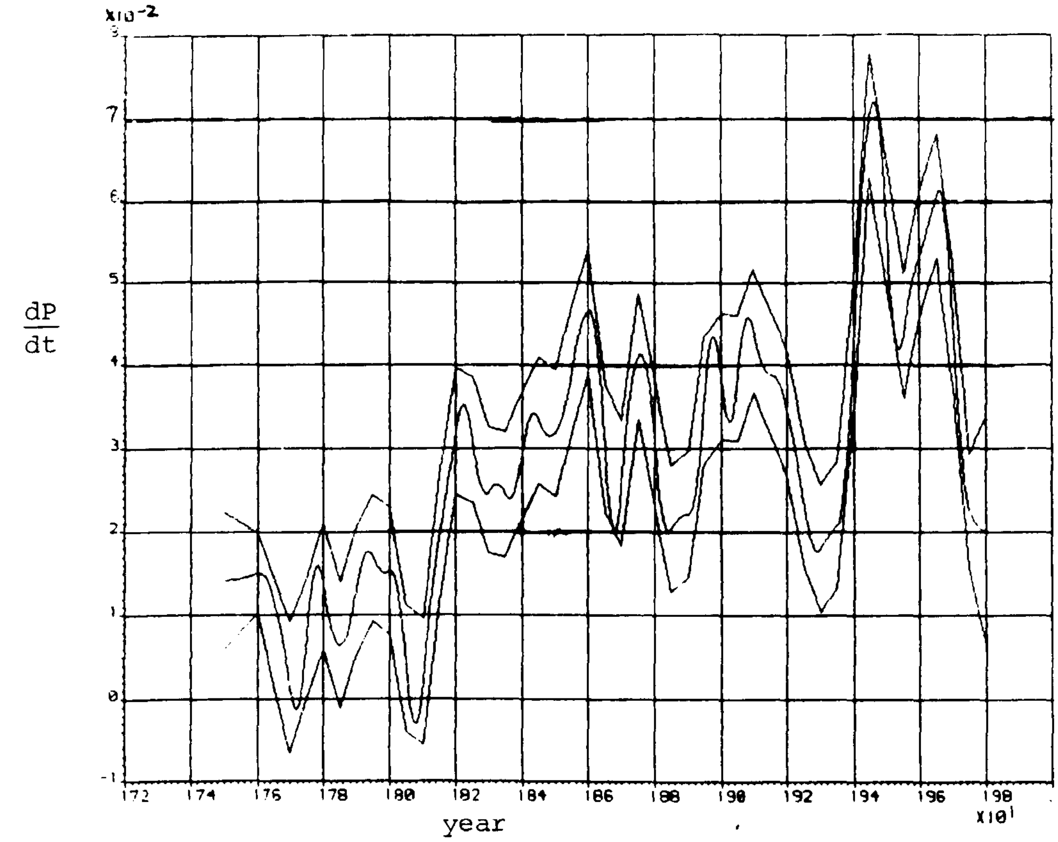
The four parameter logistic (failed to converge) - the best estimates of parameters obtained are:

$$N_\infty = 11.42,\ N_0 = 0.1550,\ c = 0.0271,\ d = -1.637.$$

Table 2.7.1.4: Percentage error of forecasts made in 1930 for Sweden

Year	Observed	3 Parameter Logistic	4 Parameter Logistic	3 Parameter Gompertz	4 Parameter Gompertz	Modified Exponential	Geometric	Linear	Quadratic	Cubic
1935	6.251	1.284	0.880	2.032	0.900	4.574	8.729	0.100	2.775	0.908
1940	6.372	1.783	1.235	2.816	1.306	5.963	11.048	0.695	3.802	1.261
1945	6.674	-0.564	-1.252	0.755	-1.117	4.457	10.381	-1.479	1.980	-1.238
1950	7.041	-3.660	-4.485	-2.050	-4.276	2.195	8.929	-4.356	-0.582	-4.497
1955	7.235	-4.269	-5.256	-2.304	-4.956	2.616	10.365	-4.723	-0.541	-5.309
1960	7.495	-5.740	-6.885	-3.415	-6.485	2.171	10.916	-5.907	-1.350	-6.999
1965	7.767	-7.311	-8.616	-4.611	-8.105	1.660	11.431	-7.154	-2.230	-8.814
1970	8.077	-9.261	-10.719	-6.183	-10.092	0.770	11.559	-8.749	-3.482	-11.027
1975	8.208	-9.183	-10.828	-5.639	-10.060	2.186	14.290	-8.268	-2.537	-11.283
1980	8.318	-8.934	-10.773	-4.888	-9.852	3.881	17.415	-7.570	-1.349	-11.417
Absolute Departures										
Sum of squared departure		2.370	3.317	0.922	2.885	0.558	8.034	2.084	0.285	3.561
M.A.D.		0.487	0.576	0.304	0.537	0.236	0.896	0.456	0.169	0.597

Figure 2.7.1.7: Derivative of the Swedish spline



Thus the upper asymptote is around 9.78 and the sum of squares of residuals is 0.00008.

Fitting to yearly estimates for the period before escalation (i.e. 1929-1941) suggests a logistic model. Data between 1941 and 1950 does not model easily and is ignored, assuming it represents a transient period prior to escalation.

Thus for Sweden two models should be used, one for between 1810 and 1930 and one for 1950 to 1980. However, for practical purposes, the three parameter logistic fitted over 1810 to 1980 and parameterised as:

$$N_\infty = 13.21\ (0.75)$$

$$N_0 = 1.31\ (0.033)$$

$$c = 0.0118\ (0.0004)$$

Variance of residuals = 3.6×10^{-4}

is an adequate fit and is sufficient for interpolation and as a basis for forecasting.

Sensitivity of the Swedish Logistic Models

To investigate further the logistic models fitted between 1810 and 1980, sampling of the census figures was carried out in that logistic models were fitted to every second census count from 1810 and compared to logistic models fitted to every second census count from 1815. The results are presented in Table 2.7.1.5.

Table 2.7.1.5

	Period of Fit	Parameters				Sum of sqs.of res.	Upper asymp-tote
		N_∞	N_0	c	d		
3 Parameter Logistic	1810-1940	9.45 (0.48)	1.16 (0.04)	0.0144 (0.0007)		0.0016	9.45 (0.48)
	1815-1940	9.57 (0.52)	1.16 (0.05)	0.0143 (0.0007)		0.0019	9.57 (0.52)
4 Parameter Logistic	1810-1940	6.93 (0.77)	0.46 (0.18)	0.0204 (0.003)	1.02 (0.32)	0.0011	7.95 (1.09)
	1815-1945	12.96 (6.94)	2.27 (2.30)	0.010 (0.006)	-1.35 (2.63)	0.0018	11.61 (9.57)
3 Parameter Logistic	1810-1980	13.29 (1.11)	1.30 (0.05)	0.0117 (0.0006)		0.0064	13.29 (1.11)
	1815-1975	13.08 (1.25)	1.31 (0.06)	0.118 (0.0007)		0.0062	13.08 (1.25)
4 Parameter Logistic	1810-1980	199.0 (4435)	6.7 (15.4)	0.0034 (0.0096)	-5.88 (15.70)	0.0047	194.3
	1815-1975	228.0 Failed to converge	7.78	0.0031	-6.95	0.0037	221.1

t* = year - 1750

This indicates that the simpler 3 parameter logistic model is more stable than the 4 parameter form, as the parameter estimates are similar between periods starting 1810 and those starting 1815 and also with estimates from unsampled fits. This finding applies for both the 1810-1945 and 1810-1980 periods.

The four parameter logistic model appears unstable, as is demonstrated by the unrealistic parameter estimates, high standard errors and convergence problems for the 1810 to 1980 period. For the period 1810-1945 this brief analysis suggests that the parameters of the four-parameter logistic can take a wide range, i.e.

	Possible range
Upper asymptote	7.95 ± 2.2m ------→ 11.61 ± 19.04m
c	0.01 ± 0.012 -----→ 0.0204 ± 0.006m

This suggests that, for sparse data, the three parameter logistic model would be a more reliable descriptor of population growth.

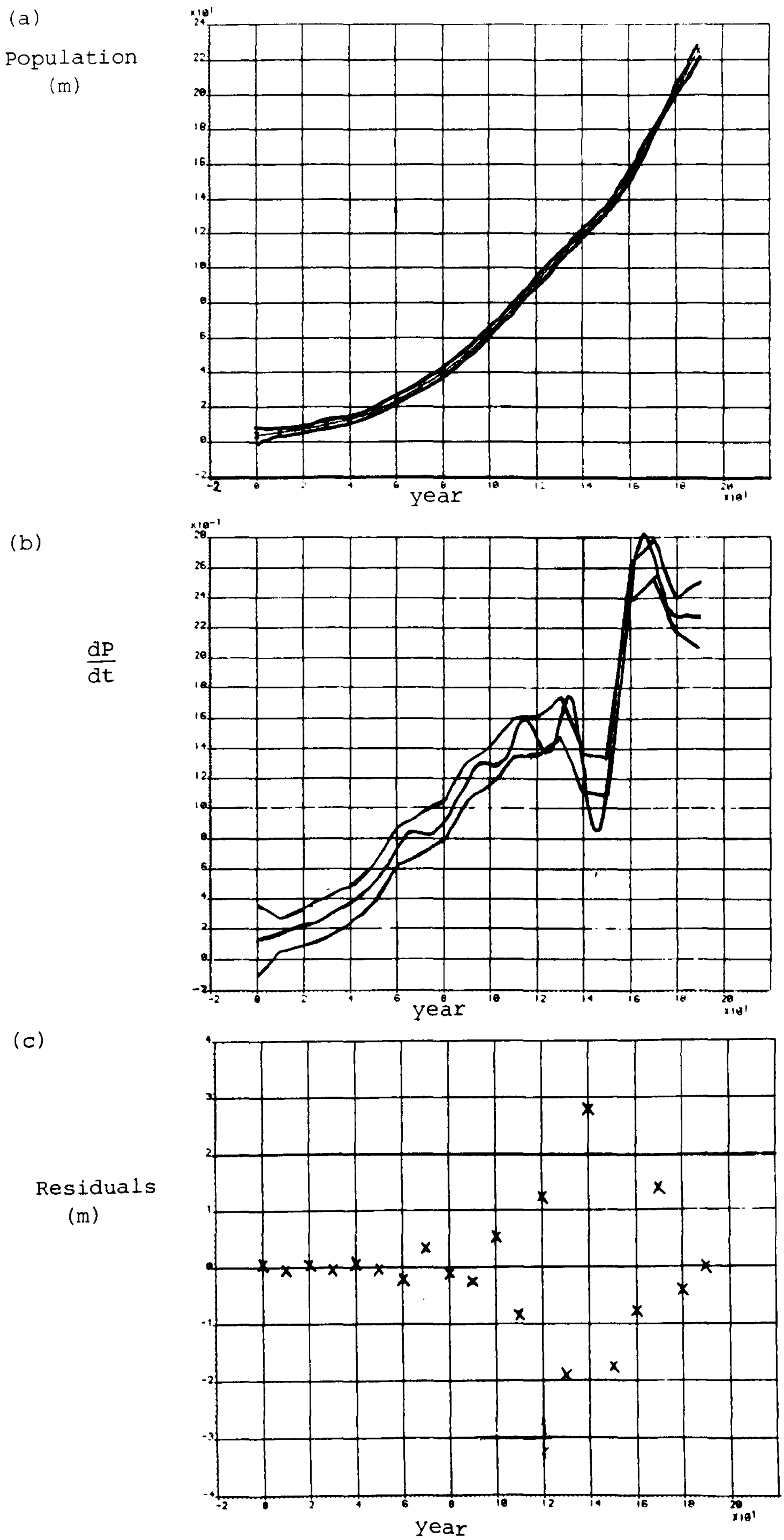
2.7.1c U.S.A.

Decennial census records of the population of the U.S.A. began in 1790 and are available for an unbroken series up to 1980 giving twenty observations. In 1959 the land mass and population was increased by the inclusion of the new states of Hawaii and Alaska - to allow for this only the populations of these states will be excluded from the analysis. It should be noted, as pointed out in Chapter 1, that the USA should not be treated as typical of Western population growth as much of the USA population growth has been determined by migration.

Exploratory Analysis

1. Spline Analysis - The spline with 95% confidence limits, its derivative and residuals are displayed in Figure 2.7.1.8. Examining the derivative (Figure 2.7.1.8b) indicates that there has either been two distinct phases of growth or 1930, 1940 and 1950 census counts have been underestimates. Examination of the birthrate minus death rate confirms that it was increased growth. Therefore escalation of models after 1940 would appear to be required. Considering the period to 1940 the spline analysis would suggest a point of inflection at 1911 giving an upper asymnptote of around 180m. For the period after 1940 one can speculate from Figure 2.7.1.8b that the point of inflection is around 1960 which gives an upper aysmptote of approximately 240m ($N_{1930} + 2(N_{1960} - N_{1930}) = 120 + 2(180-120)$). Thus, although the escalated phase of growth appears significantly to increase the upper asymptote, the escalated phase would appear to be short-

Figure 2.7.1.8: Spline - U.S.A.



lived and the renewed growth cycle may terminate before 2000 A.D.

The variance of the residuals (Figure 2.7.1.8c) show a tendency to increase suggesting the need for transformations or weighting.

2. Cusum - Figure 2.7.1.9 - does not indicate any need for escalation after 1940.

3. Time Series Representation - The regression model obtained is:

$$\frac{1}{N_t} = 1.39\frac{1}{N_{t-1}} - 0.478\frac{1}{N_{t-2}} + \epsilon_t$$

(0.192) (0.142) (standard errors in brackets)

The model is of the correct form for a logistic model and the coefficient of $-1/N_{t-2}$ should approximate to 1.39 - 1 rather than 0.478 (indicaton of a poor fit is given by the funnel-shaped residual plot in Figure 2.7.1.10).

4. Window Analysis - Results from forward and backward moving windows are displayed in Figure 2.7.1.11. The rapid increase in the upper asymptote and sum of squares of the residuals after 1950 for the forward moving window indicates the requirement for escalation and supports Leach's (1981) conclusion that escalation is required for the U.S.A.

Models Compared

A summary of simple trend models fitted between 1790 and 1930 using PTRENDS with weighting are presented in Table 2.7.1.6.

Surveying these results show that of these models the logistic models give the most suitable description of the population of the U.S.A. between 1790 and 1930. The four parameter logistic model is similar to the three parameter logistic models (the displacement parameter not being statistically significant).

Figure 2.7.1.9: Cusum of USA growth rates

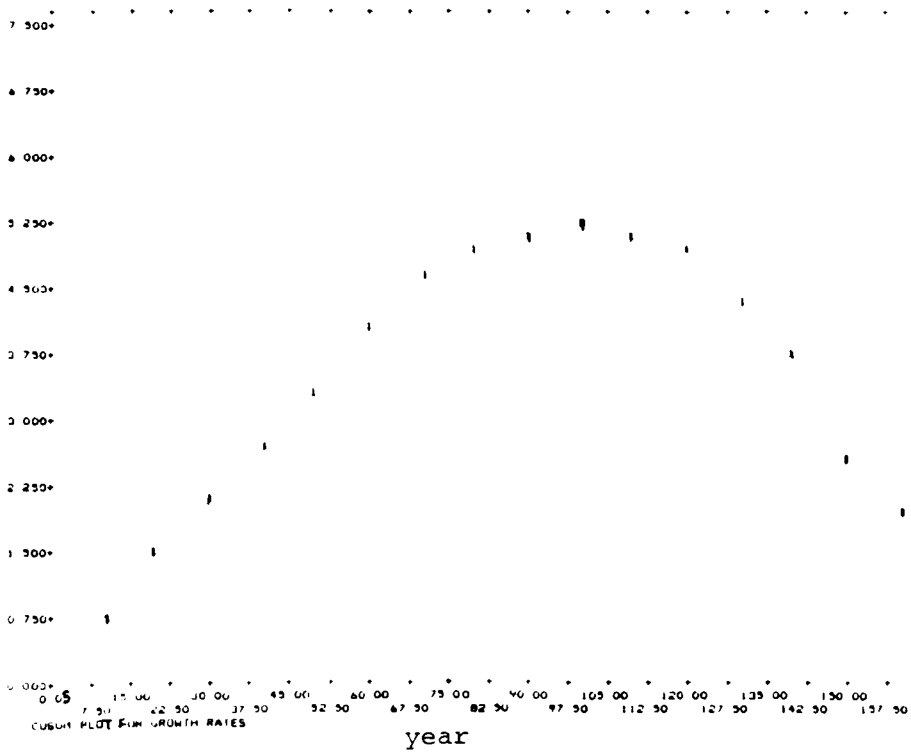


Figure 2.7.1.10: Residuals from 'linear' investigation

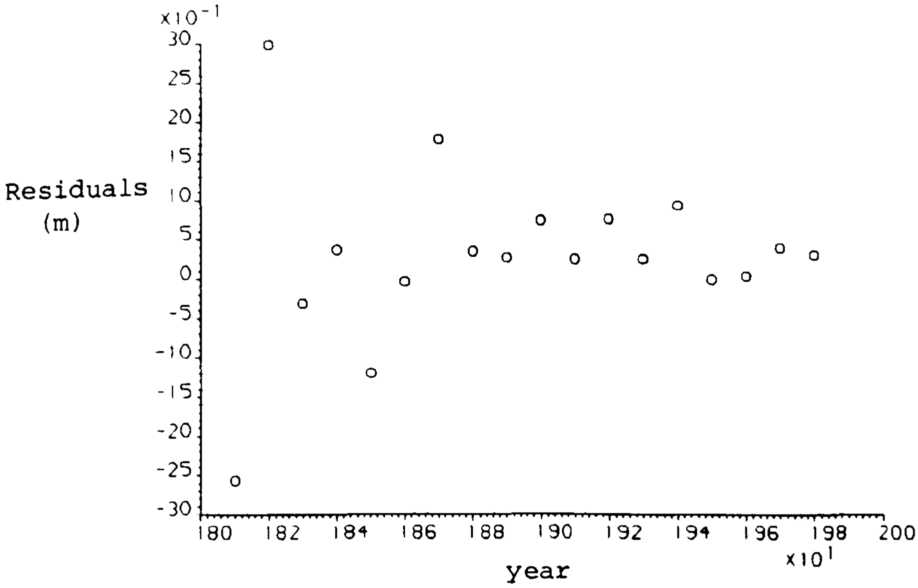
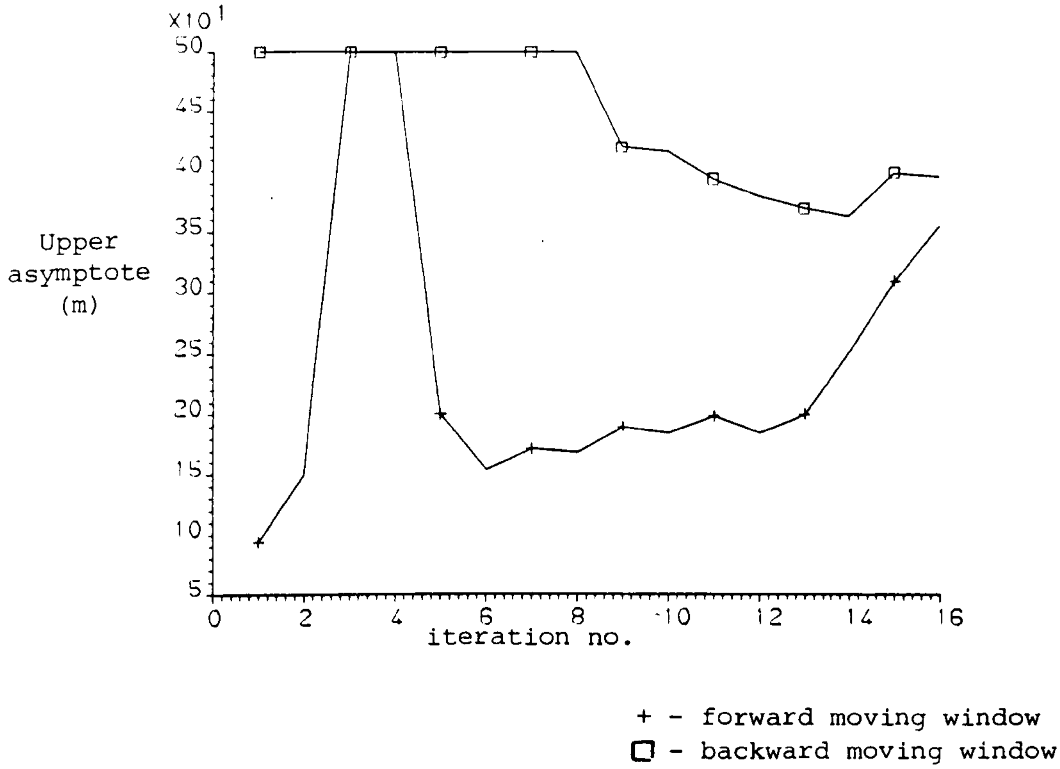


Figure 2.7.1.11a: Window analysis for the USA - 3-parameter logistic



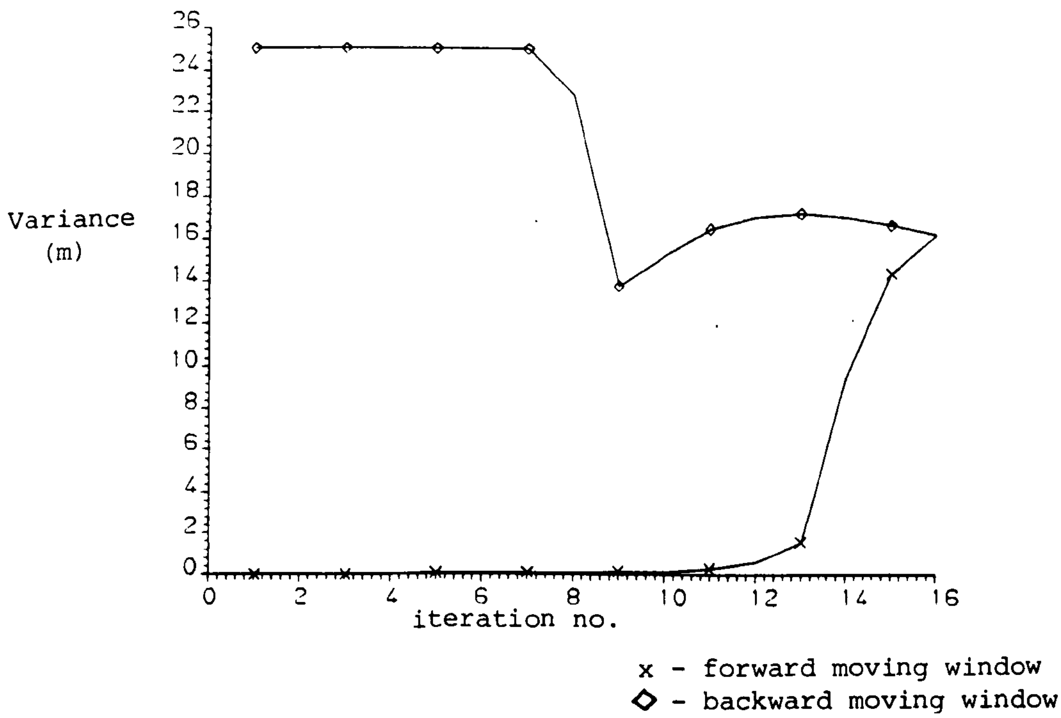
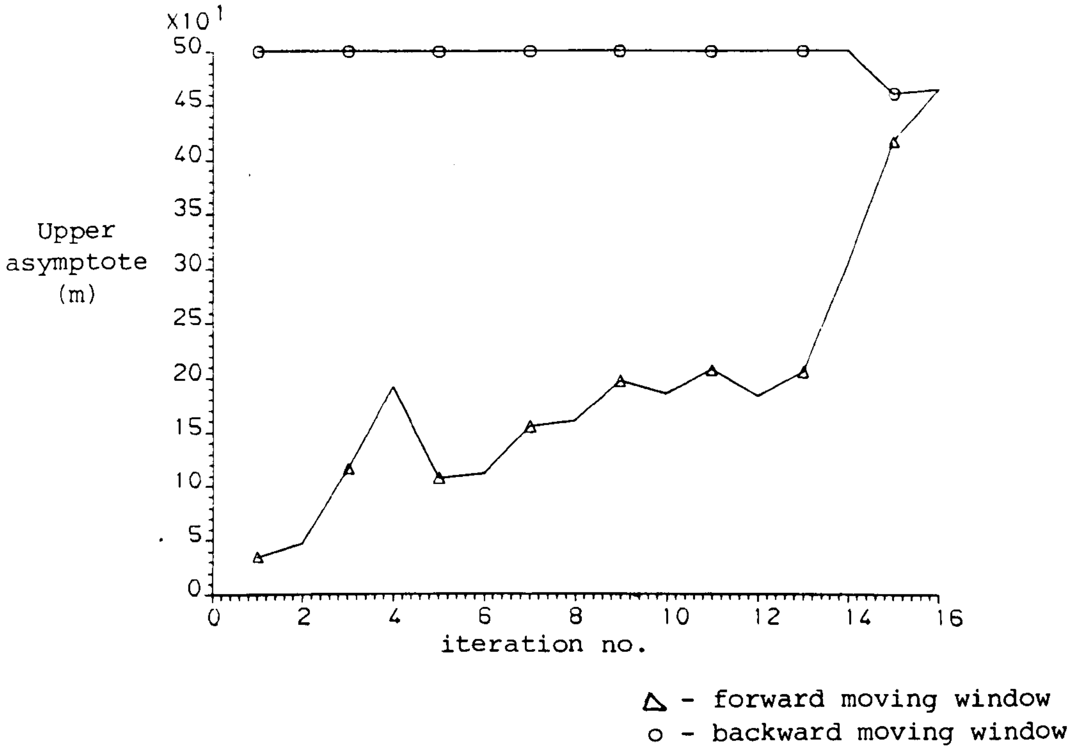


Figure 2.7.1.11b: Window analysis for the USA - 4-parameter logistic



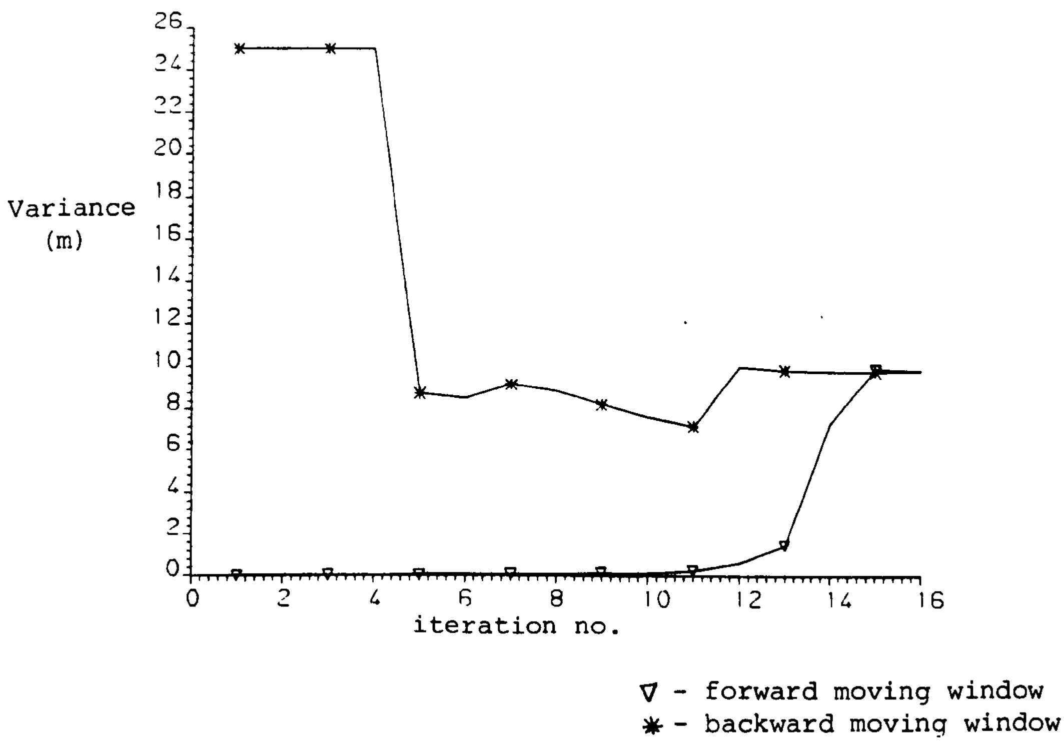


Table 2.7.1.6: Models of the Population of U.S.A. 1790-1930

for t* = year - 1790

Model	Parameter	Standard errors	Variance of fits	Sum of squares of departures SSR	Indexed Costed RSS
3 Parameter Logistic	N_∞ = 113.3 N_0 = 3.895 c = 0.032	4.87 0.03200 0.00021	0.00017 (0.00006)	4.985	1.0
4 Parameter Logistic	N_∞ = 186.15 N_0 = 3.71 c = 0.032 d = 0.23	6.522 0.14 0.00055 0.17	0.00015 (0.00006)	6.750	2.03
3 Parameter Gompertz	N_∞ = 2213.31 b = -6.421 c = 0.0057	Failed to converge	0.00164 -	-	-
Modified Exponential	A = -4.104 B = 7.778 C = 0.021	0.903 0.79 0.00088	0.00434 (0.00158)	383.035	76.92
Geometric	A = 4.616 B = 0.025	0.258 0.00068	0.01287 ±0.00470	1640.454	164.54
Linear	A = 2.94 B = 0.499	1.07 0.061	0.149 (0.05421)	5683.389	570.05
Quadratic	A = 47.54 B = 1.621 C = 0.006	Failed to converge	1.15741 -	-	-
Cubic	A = 3.987 B = 0.0942 C = 0.0025 D = 0.000022	0.095 0.0134 0.00040 0.000003	0.00061 (0.00022)	45.027	4.02

Thus the model:

$$N_t = \frac{193.3}{1 + \left[\frac{193.8}{3.9} - 1\right] e^{-0.032t*}}$$

$t* = \text{year} - 1790$

should be used to describe the population of the U.S.A. between 1790 and 1930.

Predictions made from these models for a period of fifty years up to 1980 are given in Table 2.7.1.7. The predictions for 1940 and 1950 are good (less than 5% error) but badly underestimate for periods after 1950. Of the other functions only the Gompertz models perform well but this was not suggested a priori by considering how well the models fitted. Also by 1970 the population has grown larger than the upper asymptote. This confirms the need for escalation or recourse to more complex models.

The models were refitted to the period 1790 to 1950 - these may be compared with reference to Table 2.7.1.8. Again the three parameter logistic is the most suitable model. The incorporation of the 1940 and 1950 counts has caused very little change in the parameter estimates (although the estimate of the upper asymptote is better defined for the longer period as the standard error is reduced).

Modelling the period 1960 to 1980 is difficult due to the lack of census data, and the four parameter models were not fitted. The three parameter logistic model is a good model; this is confirmed when the three parameter logistic is fitted to yearly estimates over the period 1953 to 1980 (28 points). This gives the three parameter logistic model with parameters:

Table 2.7.1.7: Percentage Error of Forecasts made in 1930 for U.S.A.

Year	Observed	3 Parameter Logistic	4 Parameter Logistic	3 Parameter Gompertz	4 Parameter Gompertz	Modified Exponential	Geometric	Linear	Quadratic	Cubic
1940	131.669	2.704	1.763	10.395	8.790	31.078	54.255	-40.889	-56.724	14.326
1950	150.697	-2.328	-3.626	12.219	8.015	41.679	73.454	-45.040	-64.104	16.421
1960	178.464	-11.837	-13.352	9.321	2.526	47.872	88.496	-50.794	-71.996	13.752
1970	202.229	-18.075	-19.771	10.419	0.694	61.184	114.080	-54.107	-77.932	15.393
1980	225.166	-23.461	-25.278	12.657	-0.269	78.709	147.447	-56.565	-83.103	18.415
Absolute Departures										
Sum of squares of Residuals		4597.9	5441.3	2059.2	302.5	59637.3	205746.4	43917.2	91271.6	4258.9
Mean Absolute Deviation		30.325	32.989	20.294	0.537	109.213	202.853	93.720	135.109	29.185

Table 2.7.1.8: Models of the Population of U.S.A. 1790-1950

for t* = year - 1790

Model	Parameter	Standard errors	Variance of fits	Sum of squares of departures SSR	Indexed Costed RSS
3 Parameter Logistic	N_∞ = 193.79 N_0 = 3.905 c = 0.0315	3.54 0.035 0.0002	0.00023 ±0.00008	29.549	1.0
4 Parameter Logistic	N_∞ = 191.07 N_0 = 3.79 c = 0.0319 d = 0.13	4.85 0.14 0.00051 0.18	0.00022 ±0.00008	32.090	1.63
3 Parameter Gompertz	N_∞ = 1409.10 b = 5.96 c = 0.0063	407.0 0.27 0.00049	0.00246 ±0.00084	277.596	9.39
Modified Exponential	A = -6.42 B = 9.96 C = 0.0183	1.42 1.26 0.001	0.00907 ±0.00311	1477.606	50.00
Geometric	A = 5.050 B = 0.023	0.381 0.00080	0.02638 ±0.00905	6393.059	108.17
Linear	A = 19.507 B = 0.9385	3.217 0.208	6.90399	-	-
Quadratic	A = 31.9665 B = 1.4344 C = 0.00305	4.023 (0.001)	0.62949	-	-
Cubic	A = 4.052 B = 0.067 C = 0.00368 D = 0.00001	0.154 0.021 0.00055 0.000002	0.00160 ±0.00055	239.412	12.15

Parameters	Standard Errors
270.8	4.48
0.19	0.579
0.0467	0.00199

with variance of residuals of $<10^{-4}$m and sum of squared deviations of 15.1.

Thus, to model the population of the U.S.A., the model proposed is:

$$N_t = \begin{cases} \dfrac{193.3}{1 + \left[\dfrac{193.3}{3.9} - 1\right]e^{-0.032t*}} & \text{for } 0 \leq t* < 17 \\ \dfrac{270.8}{1 + \left[\dfrac{270.8}{0.19} - 1\right]e^{-0.0467t*}} & \text{for } 17 \leq t* \leq 20 \end{cases}$$

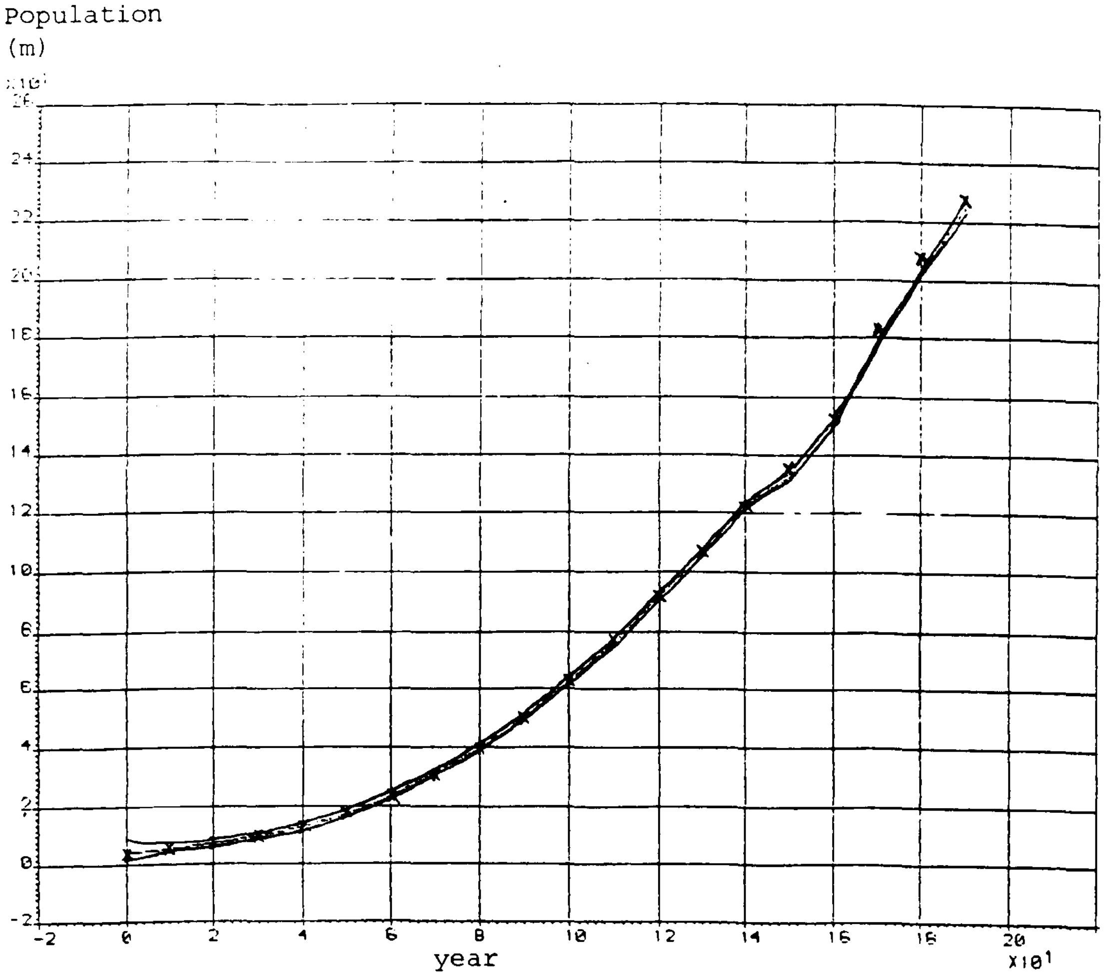
$$t* = \text{year} - 1790$$

If escalation had not been detected and the period 1790 to 1980 modelled as one unbroken series, results indicate markedly inferior fits of the models, the sum of squares of residuals greatly increasing. However, of the models the logistic models appear to give sensible parameters. The upper asymptote of the three parameter logistic model is approximately 243.6 ± 20.6m which was suggested by the spline analysis.

The upper asymptote of the four parameter (of which the parameter d is not significant) is around 266 ± 33m. These estimates of the upper asymptote are rather low as the population of the U.S.A. grew by over 10% in the period 1970 to 1980, and was 225.167m. One must question the validity of upper asymptotes as low as 250 million. Even 270m, the asymptote of the escalated model, may be low. This question will be considered in the following Chapters.

For the time being, the discontinuous model above will be used. The estimates from this model are compared to the benchmark of the spline fitted to data in Figure 2.7.1.12. This Figure illustrates that the model up to 1940 described the data extremely well as all points except the 1940 point fall within the 95% confidence intervals. After 1940 the fit is poorer, but is still good, the model tending to overestimate. (All residuals having a positive sign after 1940 does not suggest a fault with the least squares algorithm as the model was derived from different data, i.e. yearly estimates.)

Figure 2.7.1.12: Model for the USA compared to the spline



2.7.1d <u>France 1801-1982</u>

For France eighteen censuses have been taken irregularly. In the censuses between 1871 and 1918 the provinces of Alsace and Lorraine have not been included in the census counts. Estimates of the population of these regions have been added to the appropriate census values.

The exploratory analysis of France indicates that:

(i) compared to populations discussed the 95% confidence intervals on the spline are wider indicating that one should anticipate less precise fits of the models;

(ii) there appears to be escalated growth after 1950.

The presence of two cycles of growth is confirmed by examining the derivative of the spline and window analysis.

Comparing the models and reviewing predictions (see Table 2.7.1.9) made in 1930 further highlights the need for escalation after 1954.

After examining different models, the best strategy to model the French population is to use logistic models for the two growth phases. These models are parameterised as follows:

Period 1801 to 1950 Parameter	Standard error	Period 1951 to 1982 Parameter	Standard error
N_∞ 41.59	0.97	N_∞ 57.283	0.593
N_0 27.22	0.57	N_0 0.002	0.002
α 0.0216	0.0033	α 0.0732	0.0064
Variance = 5.89×10^{-4}		Variance = 2.725×10^{-3}	

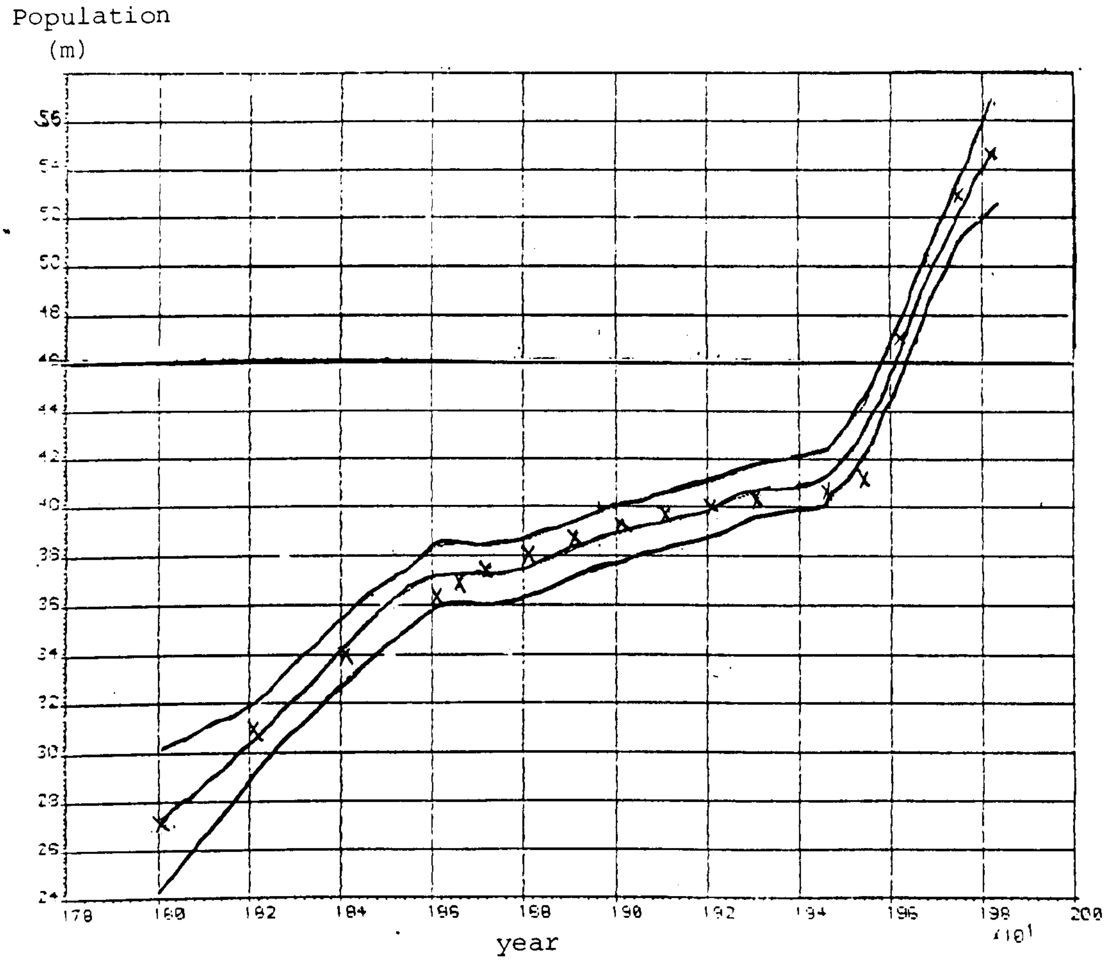
$t^* = \text{year} - 1801$

As can be observed from Figure 2.7.1.13, the above "two phase" model compares favourably to the benchmark of the spline. Hence this model is used to describe the growth of the population of France.

Table 2.7.1.9: Percentage Error of Forecasts made in 1931 for France

Year	Observed	3 Parameter Logistic	4 Parameter Logistic	3 Parameter Gompertz	4 Parameter Gompertz	Modified Exponential	Geometric	Linear	Quadratic	Cubic
1946	40.507	0.47	0.84	0.88	0.63	8.54	9.85	8.38	-0.60	72.94
1954	42.777	-4.49	-4.10	-4.05	-4.34	4.75	6.49	4.58	-6.16	73.33
1962	46.520	-11.89	-11.48	-11.43	-11.75	-1.86	0.15	-2.05	-14.15	68.88
1968	49.779	-17.49	-17.07	-17.01	-17.35	-7.01	-4.78	-7.21	-20.20	64.94
1975	52.656	-21.83	-21.40	-21.34	-21.70	-10.68	-8.16	-10.90	-25.16	64.26
1981	53.962	-23.60	-23.15	-23.08	-23.46	-11.66	-8.82	-11.90	-27.56	67.68
1982	54.335	-24.10	-23.66	-23.59	-23.97	-12.07	-9.19	-12.31	-28.16	67.79
Absolute Deviations										
Sum of squares of residuals		575.84	552.28	548.66	568.34	143.25	95.26	148.00	782.33	776.39
M.A.D.		7.67	7.51	7.48	7.62	4.05	3.32	4.11	8.99	33.21

Figure 2.7.1.13: The model for France compared to the spline



x - fitted point

2.7.1e Japan 1852-1980

Twelve censuses in Japan have been taken irregularly between 1852 and 1980.

Exploring the Data

1. Spline Analysis - the spline and its 95% confidence intervals are given in Figure 2.7.1.14. The derivative of this indicates only half the inverted '∩', see Figure 2.7.1.14b, and there is indication that the maximum rate of growth is at 1980 which would give an upper asymptote of 230m for the three parameter logistic.

The prospect of this is questionable as current trends in Japan indicate a balance between births and deaths and various projections made by the Japanese government suggest (see the 1984 Japanese Yearbook) that, after a transitional phase for age structure effects to work their way out, the Japanese population will stabilise around 120m by the beginning of the twenty-first century.

To investigate this further, the spline analysis was performed on the yearly estimates of the Japanese population over the period 1872 to 1983. The derivative of this spline (see Figure 2.7.15b) gives a much more complex display; one interpretation of this is that there have been three phases of growth. The phases of growth could be from 1870 to 1938, from 1940 to 1958 and 1960 to 1983. This suggests the need for three escalated sigmoids, or a more complex model.

2. Window Analysis - was carried out for the census counts and yearly estimates. This analysis indicates that there is possibly a need to split the modelling into two parts, from 1852 to 1930, and 1930 to 1982.

From fits of the simple models to the census counts of the

Figure 2.7.1.14: The spline for Japan (census)

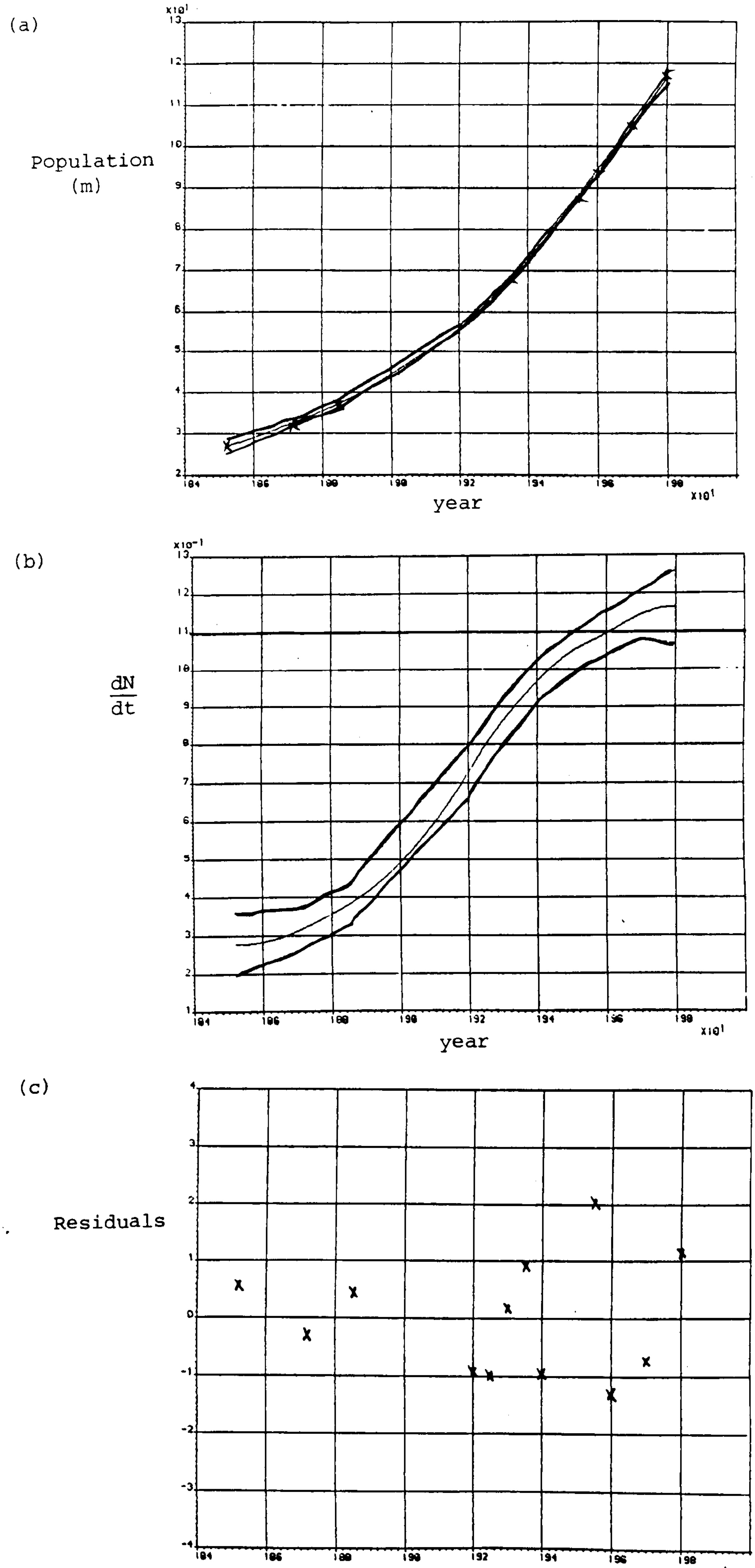
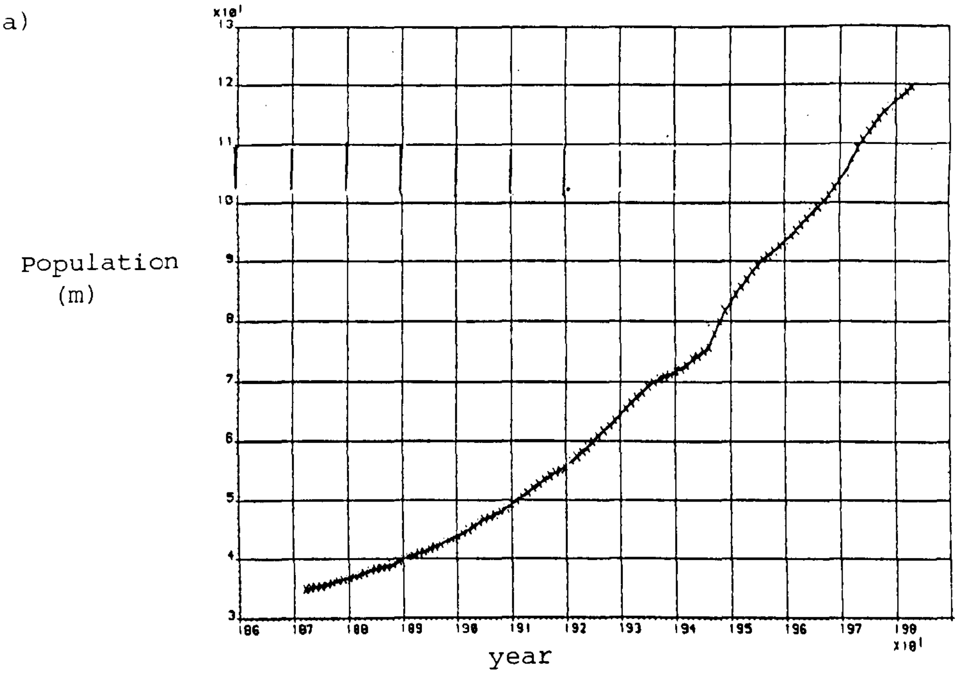
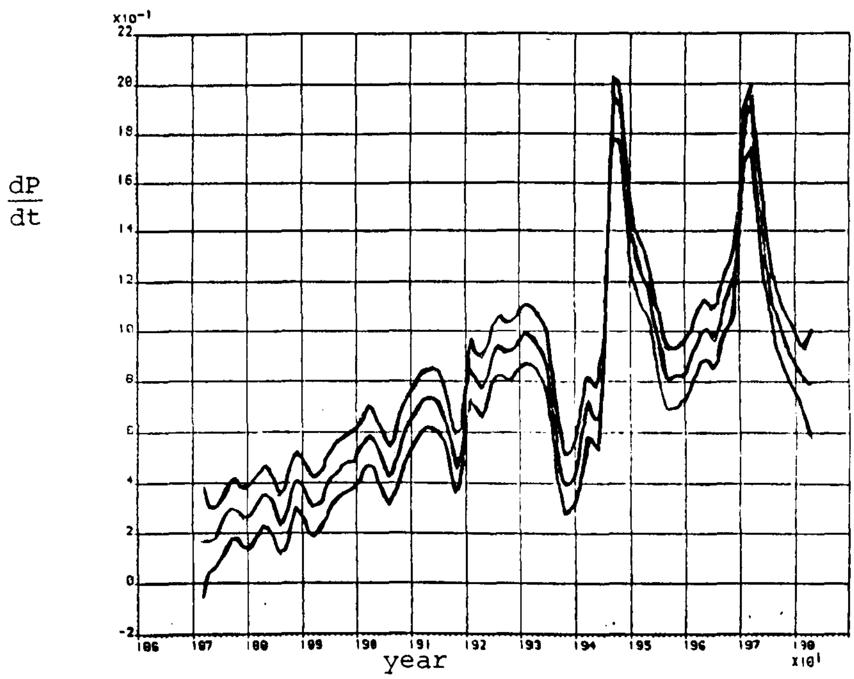


Figure 2.7.1.15: The spline for Japan (yearly estimates)

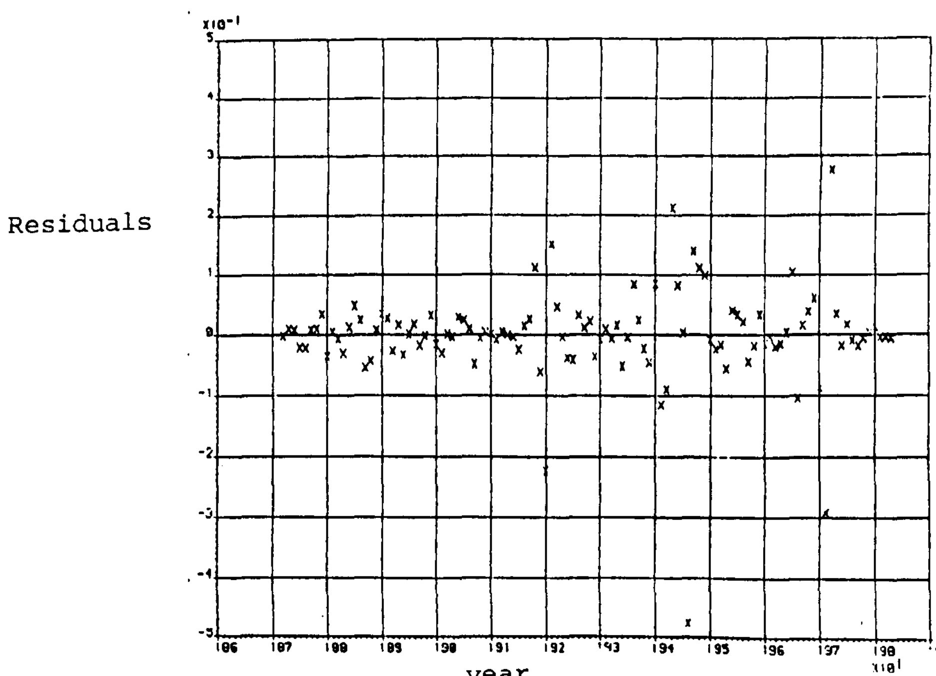
(a)



(b)



(c)



Japanese population, the four parameter logistic gives the 'best' description of the Japanese population over this period. However, the exploratory analysis of the population indicated that one continuous simple model was inadequate, and this is supported by the high upper asymptote of around 240 million, given by the four parameter logistic. This is considered high as Japan, whose 1980 population number of around 117 million, is one of the world's most densely populated regions with a density of over 300 people per square kilometre and is displaying a tendency to replacement (or less) fertility levels, and this suggests that there is little prospect for the almost doubling of population in a little over one hundred years as the four parameter logistic predicts.

Splitting the period and using yearly estimates as suggested by exploring the data gives the 4-parameter logistic model:

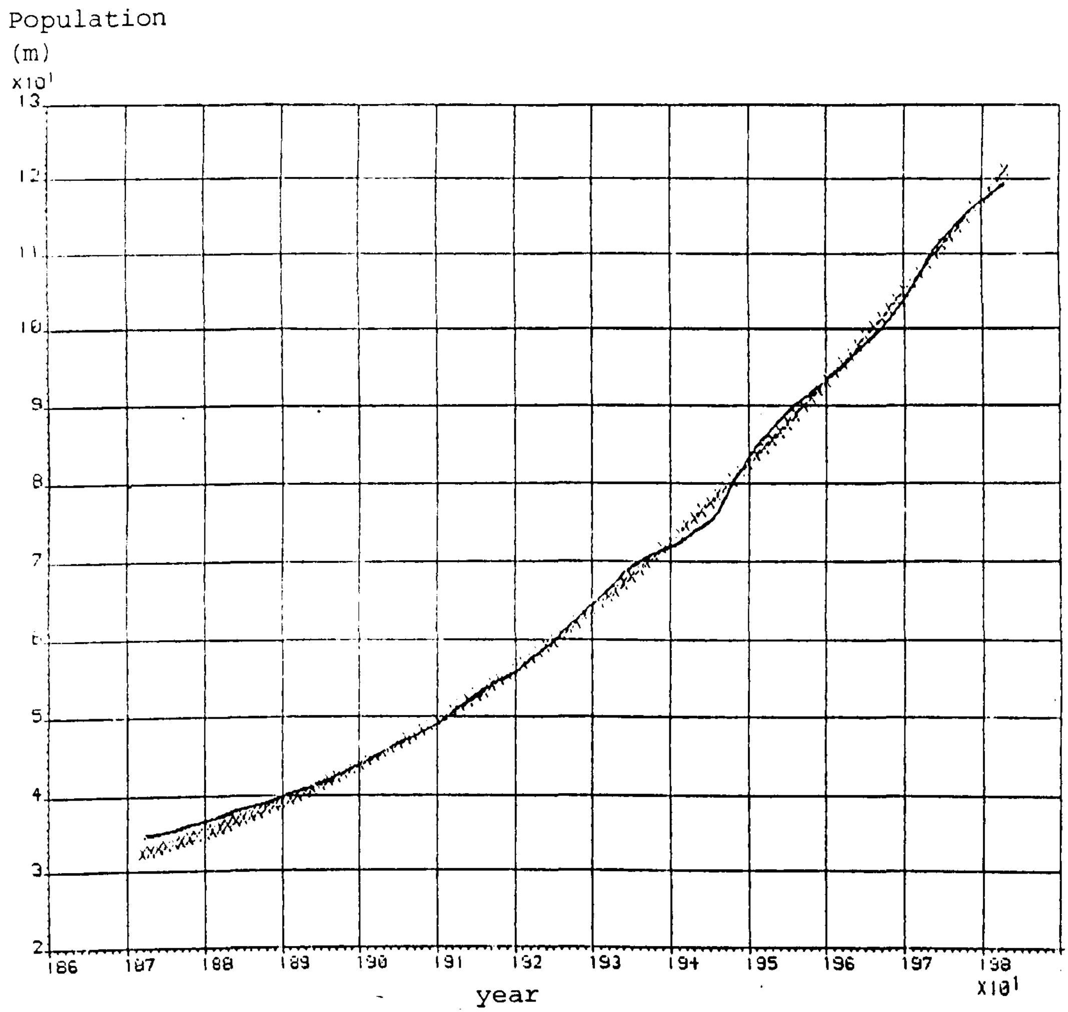
$$N_t = \begin{cases} \dfrac{32.04}{1 + \left[\dfrac{32.04}{0.0074} - 1\right]e^{-0.076t*}} + 33.77 & \text{for the period 1872 to 1919} \\[2ex] \dfrac{70.89}{1 + \left[\dfrac{70.89}{0.0005} - 1\right]e^{-0.073t*}} + 56.38 & \text{for the period 1922 to 1983} \end{cases}$$

which compares favourably with the benchmark of the spline, as can be observed from Figure 2.7.1.16. This confirms the belief that the model fitted to the census counts gave too high an upper asymptote.

2.7.1f Other Developed Nations

A summary of fits of the simple trend models to various other developed nations is given in Table 2.7.1.10. This Table displays only the "best" model out of the group of models considered in terms of closeness of fit, sensible parameters,

Figure 2.7.1.16: The model for Japan compared to the spline



x - fitted point

interpolation ability, fitting subsections of the data and extrapolation compared to observations. The logistic based models are the most suitable models for 18 out of the 22 populations that are considered in Table 2.7.1.10 and all but one are sigmoid models.

2.7.1 Conclusion

The review of the developed nations gives support to the view that population tends to develop sigmoidally. The logistic model appears to fit in most cases very well, especially the "mature" Western populations, since these illustrate the principal features of a sigmoid (i.e. rapid growth followed by growth slowing and moving to zero growth).

However, some populations, notably the USA and France, require escalation which supports Pearl's (1925) view of population growing in cycles (or 'epochs'). Early identifications of these changes are vital for sensible forecasting and this point will be considered at length in later Chapters.

In the next two Sections, examples are given of populations which do not show such well-defined sigmoid patterns.

Table 2.7.1.10: Summary of Best Models of Other Developed Nations

Nation	Model	Parameter	Standard Errors	Variance
Australia + 1861-1981 13 counts	3 Parameter Logistic t*= yr-1801	N_∞ = 20.67 N_0 = 0.26 c = 0.0277	3.63 0.04 0.002	5.6×10^{-3}
Austria 1910-1981 7 counts	4 Parameter Logistic t*= yr-1801	N_∞ = 1.30 N_0 = 0.01 c = 0.079 d = 6.59	0.87 0.03 0.063 0.12	2.6×10^{-4}
Belgium 1816-1980 15 counts	4 Parameter Gompertz t*= yr-1801	N_∞ = 6.50 N_0 = -9.44 c = 0.0231 d = 4.09	0.39 1.61 0.0022 0.07	7.6×10^{-4}
Bulgaria 1881-1975 13 counts	3 Parameter Gompertz t*= yr-1851	N_∞ = 11.34 N_0 = 2.85 c = 0.0187	1.73 0.17 0.0037	4.3×10^{-3}
Canada 1861-1941 10 counts	4 Parameter Logistic t*= yr-1851	N_∞ = 13.19 N_0 = 0.03 c = 0.049 d = 2.81	4.55 0.25 0.014 0.43	1.24×10^{-3}
Canada 1951-1981 7 counts	3 Parameter Logistic t*= yr-1851	N_∞ = 28.45 N_0 = 0.004 c = 0.0594	0.49 0.0012 0.0023	7.8×10^{-5}
Denmark 1769-1980 26 counts	4 Parameter Logistic t*= yr-1750	N_∞ = 6.96 N_0 = 0.10 c = 0.0212 d = 0.64	0.50 0.014 0.001 0.027	4.5×10^{-4}
Finland 1750-1980 24 counts	3 Parameter Logistic t*= yr-1750	N_∞ = 7.59 N_0 = 0.43 c = 0.0146	0.67 0.011 0.0005	2.13×10^{-3}
Greece 1838-1981 18 counts	4 Parameter Logistic t*= yr-1801	N_∞ = 10.14 N_0 = 0.04 c = 0.0428 d = 0.66	1.32 0.02 0.0059 0.12	9.67×10^{-3}
Hungary 1910-1980 8 counts	3 Parameter Gompertz t*= yr-1900	N_∞ = 14.84 b = -0.74 c = 0.01	4.02 0.25 0.0058	5.04×10^{-4}

+ Indications that escalation is required after 1951 but modelling to account for escalation does not greatly change the parameterisation of the model, thus the model is left in its three parameter form.

Nation	Model	Parameter	St. Errors	Variance
Iceland 1767-1980 17 counts	4 Parameter Logistic t*= yr-1750	N_∞ = 0.85 N_0 = 0.0007 c = 0.0255 d = 0.05	0.86 0.0005 0.004 0.01	4×10^{-4}
Italy 1833-1981 18 counts	4 Parameter Logistic t*= yr-1801	N_∞ = 76.26 N_0 = 3.77 c = 0.0175 d = 15.25	14.83 1.14 0.0025 1.56	1.4×10^{-4}
Luxemburg 1839-1980 26 counts	4 Parameter Logistic t*= yr-1801	N_∞ = 0.28 N_0 = 0.015 c = 0.021 d = 0.15	0.11 0.012 0.0082 0.02	10^{-4}
New Zealand 1901-1981 12 counts	4 Parameter Logistic t*= yr-1801	N_∞ = 5.06 N_0 = 0.013 c = 0.035 d = 0.58	3.73 0.03 0.019 0.41	4×10^{-4}
New Zealand 1858-1981 19 counts	3 Parameter Gompertz t*= yr-1801	N_∞ = 3.87 b = -9.72 c = 0.0187	0.58 1.14 0.0023	4.89×10^{-3}
New Zealand North Island 1901-1981 10 counts	3 Parameter Logistic t*= yr-1801	N_∞ = 6.58 N_0 = 0.04 c = 0.0253	3.15 0.02 0.0036	3.6×10^{-4}
New Zealand South Island 1901-1981 10 counts	4 Parameter Logistic t*= yr-1801	N_∞ = 0.66 N_0 = 0.0002 c = 0.051 d = 0.38	0.21 0.0001 0.019 0.16	7×10^{-4}
Norway 1769-1980 19 counts	4 Parameter Gompertz t*= yr-1750	N_∞ = 7.18 b = 5.28 c = 0.0084 d = 0.63	1.66 0.44 0.0015 0.06	1.37×10^{-3}
Portugal 1854-1981 16 counts	4 Parameter Logistic t*= yr-1801	N_∞ = 8.19 N_0 = 0.18 c = 0.028 d = 3.15	1.40 0.13 0.0055 0.31	8.3×10^{-4}
Serbia 1834-1910 15 counts	4 Parameter Logistic t* = 1800	N_∞ = 6.10 N_0 = 0.09 c = 0.0356 d = 0.46	5.97 0.01 0.015 0.20	2.5×10^{-3}

Nation	Model	Parameter	St. Errors	Variance
Spain 1900-1981 9 counts	4 Parameter Logistic t*= yr-1750	N_∞ = 65.77 N_0 = 0.34 c = 0.0207 d = 11.55	76.53 0.78 0.0117 4.18	8.38×10^{-5}
Switzerland 1837-1980 15 counts	4 Parameter Logistic t*= yr-1800	N_∞ = 6.33 N_0 = 0.02 c = 0.0348 d = 2.24	4.55 0.06 0.0183 0.29	1.2×10^{-2}
Yugoslavia 7 counts	Geometric t*= yr-1900	A = 9.82 B = 0.0103	0.20 0.000365	3.6×10^{-4}

2.7.2 Models of the Population of the Emerging Nations

Several nations such as Brazil and some S.E. Asian nations such as North and South Korea appear to be making a transition from underdeveloped nations to developed ones (Oshima, 1983). This process has been aided in many cases by transfers of technology from developed nations. This is likely to have immense effects on the populations of these nations as they undergo associated social and cultural changes. Historically these nations have been characterised as having predominantly rural populations with fairly high levels of fertility and mortality.

One aspect of transition is the reduction in mortality and the development of large urban/industrial populations.

To find how such changes affect modelling of these populations a similar analysis is embarked on as was carried out for the developed nations, devoting specific attention to the populations of Mexico, Brazil and India. India has been rather optimistically categorised as an emerging nation but is at the other end of the "scale" from Brazil having many features of an underdeveloped nation.

2.7.2a Mexico

Mexico has exhibited both rapid economic and population growth over recent decades. To find how this has affected the models fitted, the ten census counts of Mexico, which have been taken irregularly between 1895 and 1980, will be considered.

Figure 2.7.2.1b illustrates the derivative of the spline that was fitted to the data and suggests the curve of the Mexican population has not yet encountered a point of inflection. Because there has been a major disruption to growth of the Mexican population over the period 1910 to 1921, possibly as a consequence of the disruption caused by the Mexican revolution resulting in under enumeration, high emigration to the USA and the 1918 influenza epidemic, the modelling of the Mexican population will only be considered over the period 1921 to 1980 (7 observations).

The models fitted over this period are summarised in Table 2.7.2.1. From consideration of Table 2.7.2.1 the modified exponential appears as the most suitable growth model. However, the cubic gives the closest fit and, when fitted over the period 1921 to 1970, gave the best prediction of the 1980 population size. The prediction from the cubic was very close having an error of only 0.02%; the next best was the modified exponential with an error of 4.7%. Thus, since all models are very much bettered by the cubic and no inflection in the period 1921 to 1980 has been observed, one one has to conclude that the growth models are no better than the more simple cubic form.

This will likely continue to be the case until more censuses are available of the Mexican population and a point of inflection is observed. However, it should be noted that the modified exponential is a good model and could be used to describe the population of Mexico over the period 1921 to 1980. A point of inflection in the Mexican population growth may be manifested in the near future, not as a result in changes in fertility or mortality, but as a consequence of high levels of voluntary migration to the USA.

Figure 2.7.2.1: The spline for Mexico

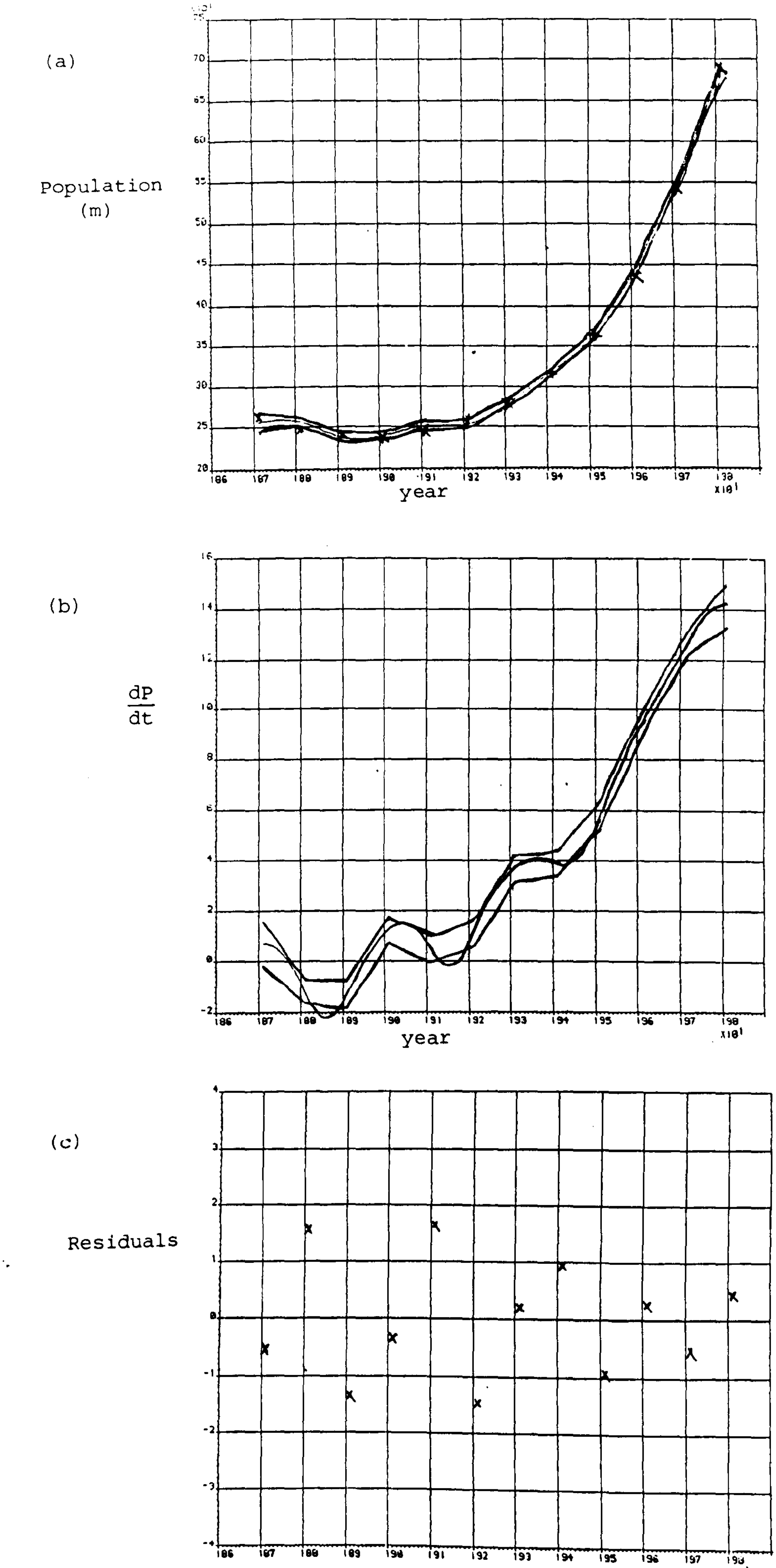


Table 2.7.2.1: Models of the Population of Mexico 1921-1980 (7 counts)

t* = year - 1880.0

Model	Parameter	Standard errors	Variance of Residuals	Sum of squares of Residuals RSS	Indexed Costed RSS
3 Parameter Logistic	N_∞ = 69013 N_0 = 1.92 c = 0.0267	45587609 1.01 0.008	8.21×10^{-3}	42.648	224.47
4 Parameter Logistic	N_∞ = -257904 N_0 = 0.22 c = 0.043 d = 9.56	Failed to converge	3.59×10^{-4}	1.249	9.86
3 Parameter Gompertz	N_∞ = 10714 b = -9.25 c = 0.0044	142590 11.64 0.010	1.33×10^{-2}	85.362	449.26
4 Parameter Gompertz	N_∞ = 4563 b = -14.56 c = 0.0092 d = 12.10	Failed to converge	1.3×10^{-4}	0.135	1.08
Modified Exponential	A = 9.56 B = 0.225 C = 0.428	0.57 0.045 0.0016	2.69×10^{-4}	1.248	6.58
Geometric	A = 1.92 B = 0.0267	0.31 0.0016	6.56×10^{-3}	42.553	111.97
Linear	A = -31.30 B = 0.606	8.53 0.101	3.86×10^{-2}	467.351	1229.87
Quadratic	A = 90.57 B = -2.11 C = 0.0147	12.14 0.27 0.0015	1.80×10^{-3}	12.146	63.92
Cubic	A = -55.86 B = 2.67 C = -0.036 D = 0.00018	10.90 0.30 0.0016 0.00007	8.89×10^{-5}	0.125	1.0

2.7.2b Brazil

Since 1950 Brazil has developed quickly economically and large urban centres have emerged. But there is a large gap between the wellbeing of the rich and that of the majority of the population; also pollution and general degradation of the environment are major problems for Brazil. There have been thirteen censuses taken irregularly in Brazil since 1808.

The spline of the population and its derivative do not support sigmoid models (see Figure 2.7.2.2) and simple exponential models would seem to be appropriate. However, the last point of the curve of Brazilian growth may indicate a turning point and hence the point of inflection which would give an upper asymptote of around 200m if the three parameter logistic is a suitable model.

Consideration of the window analysis reveals that a good model of the Brazilian population may be obtained if escalation was allowed for after 1900. (The window analysis also highlighted difficulty in setting the upper asymptote.)

From consideration of successive fits in the window analysis the optimal model appears to be:

$$N_t = \begin{cases} \dfrac{185.82}{1 + \left[\dfrac{185.82}{2.66} - 1\right] e^{-0.0197t^*}} & \text{for 1808-1900} \\[2ex] \dfrac{152.61}{1 + \left[\dfrac{152.61}{0.0016} - 1\right] e^{-0.0665t^*}} + 26.50 & \text{for 1920-1980} \end{cases}$$

$$t^* = \text{year} - 1801$$

From Figure 2.7.2.3 one observes that this model compares favourably to that of the spline.

If the whole period from 1808 to 1980 with one simple sigmoid model gives unsatisfactory results, in terms of

Figure 2.7.2.2: The spline for Brazil

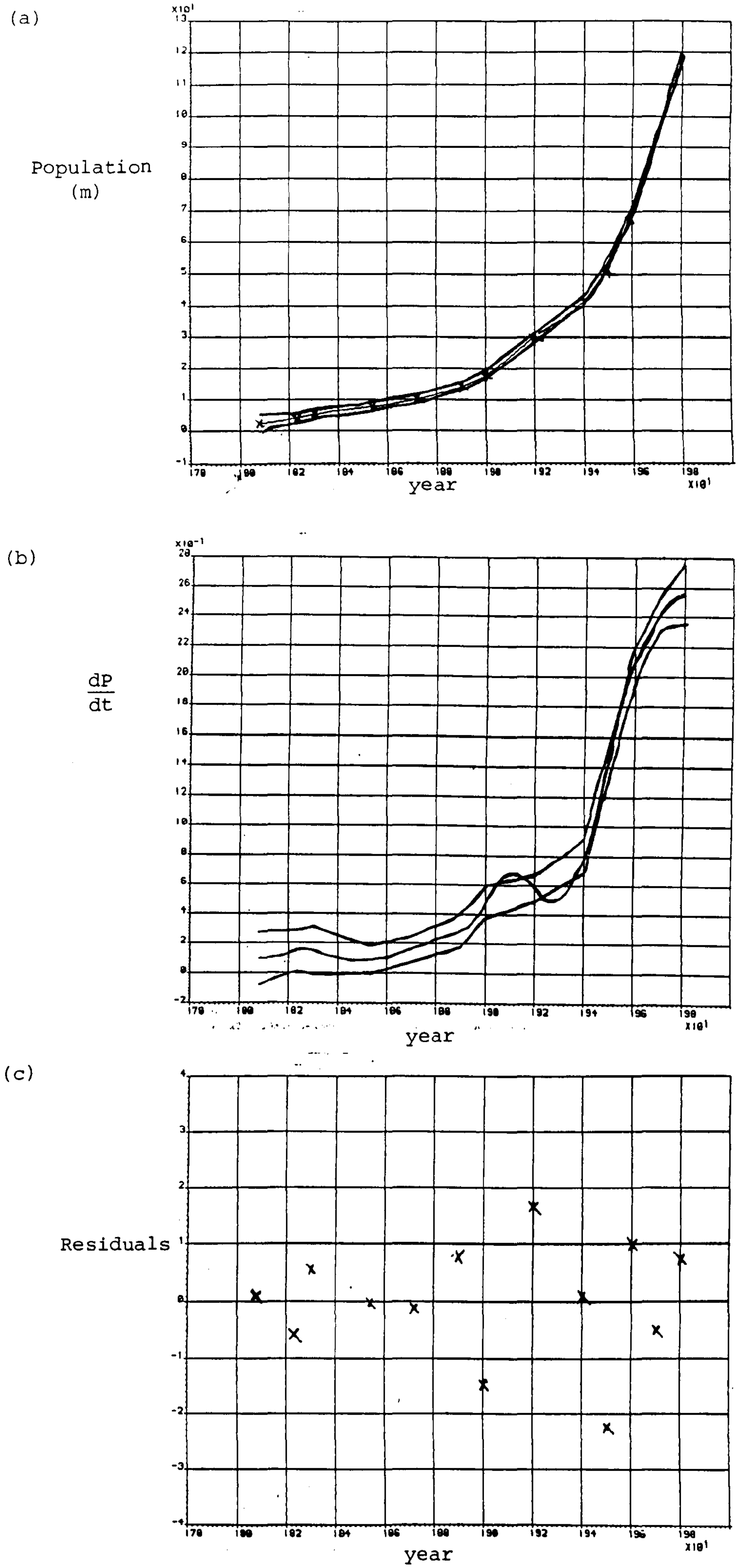
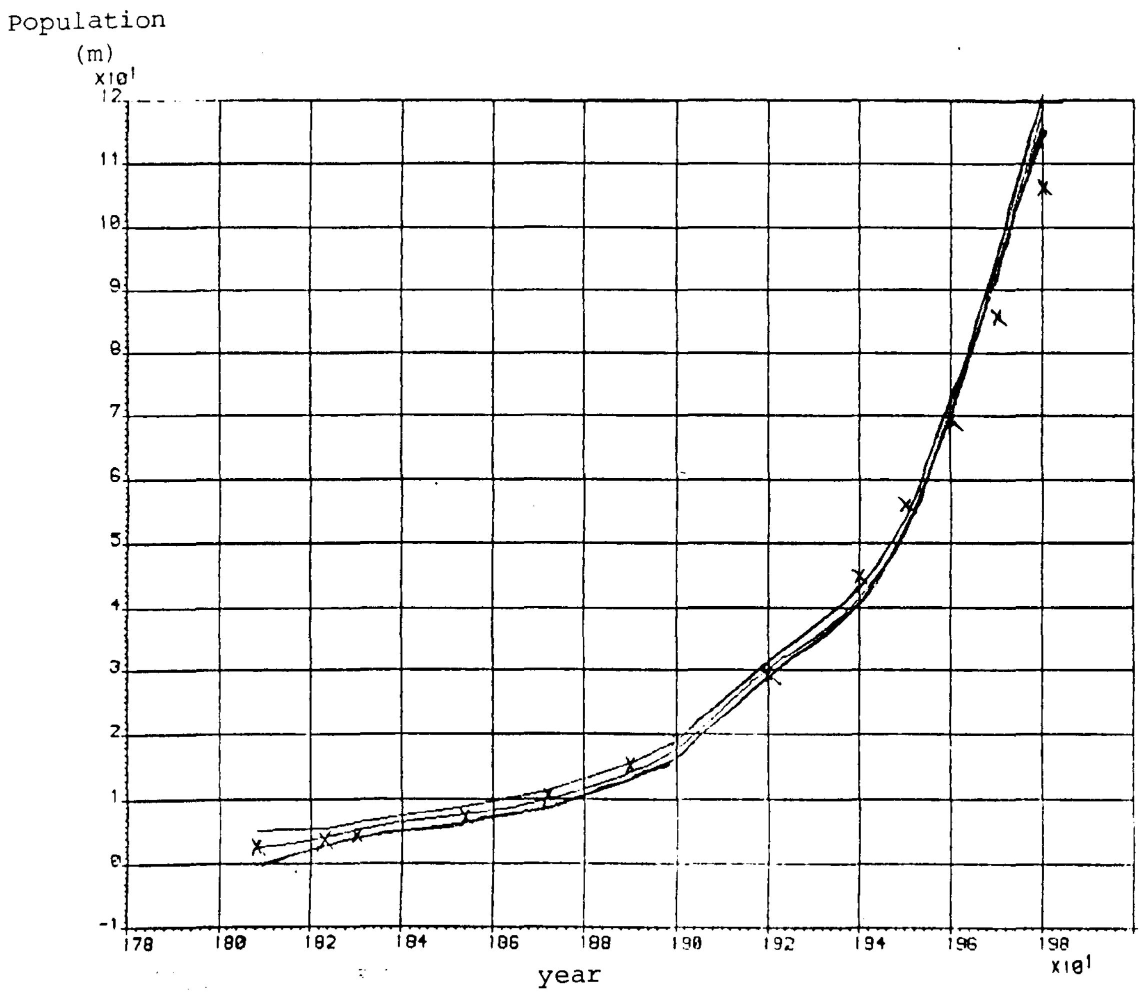


Figure 2.7.2.3: Model for Brazil compared to the spline



x - fitted value

specification of the parameters, the geometric model $N_t = 2.286e^{0.0214t^*}$, t^* = year - 1801, is easily the "best" fit.

Whether one chooses the more complex model that was discovered in the window analysis, or the simple geometric model, depends on whether one considers that the population of Brazil has followed a growth curve and passed its point of inflection. If it has not, then one would be as well using the geometric model.

Fertility in Brazil is falling, but so too is mortality as life expectancy increases. The effect of this may be a continuance of geometric growth for some time.

2.7.2c <u>India</u>

This large land mass is composed of a population of great diversity, both in economic and demographic terms, and a strong case can be made to include India in the next Section as an underdeveloped nation.

To model the population of India the nine decennial census counts between 1901 and 1981 were used. Prior to this period estimates of India's population fluctuated considerably, partly as a result of high mortality and enumeration difficulties.

The spline fitted to these counts and its derivative indicate a similar situation to Brazil.

By comparing models fitted over the period up to 1951 and extrapolations up to 1981 (see Table 2.7.2.2), one observes that no model fits nor projects the Indian population particularly well (the four parameter logistic being the best), and no model can better the cubic at predicting the 1961, 1971 and 1981 population.

When the entire period is considered the Gompertz models perform better; the four parameter Gompertz is markedly the closest fit and should give good interpolations. This model is parameterised as:

N_∞ = 11961m (16308)

N_0 = -22.57m (6.34)

c = 0.011 (0.004) t^* = year - 1801

d = 234.70m (5.35)

variance = 2×10^{-4}

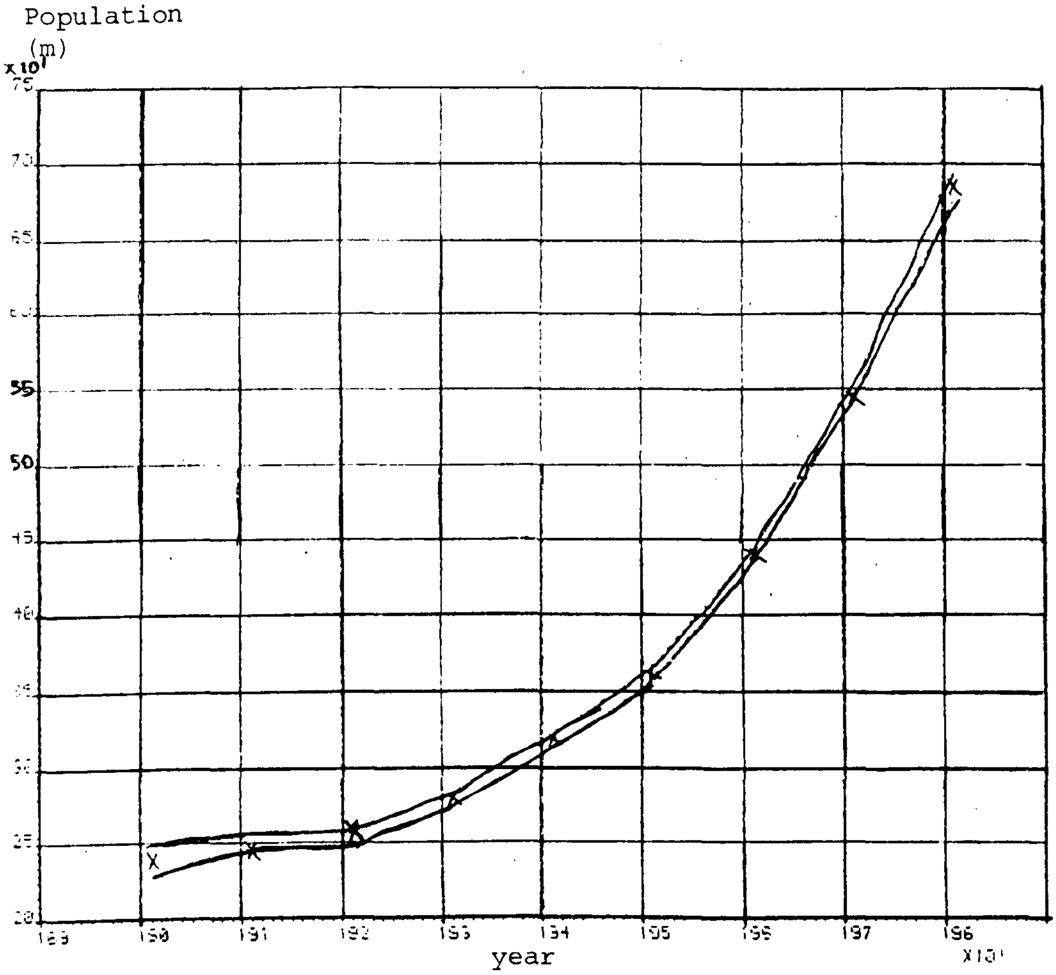
Unfortunately the specification of the upper asymptote of this model is vague; this is a consequence of the point of inflection in growth rates being encountered relatively recently.

For India, the most appropriate model is the four parameter Gompertz, which, as can be seen from Figure 2.7.2.4, compares favourably with the benchmark of the spline. However, more data is required for confident estimates of the parameters.

Table 2.7.2.2: Percentage error for fits to the Indian population up to 1951, extrapolated to 1981

Year	3 Parameter Logistic	4 Parameter Logistic	3 Parameter Gompertz	4 Parameter Gompertz	Modified Exponential	Geometric	Linear	Quadratic	Cubic
1961	-15.1	-11.7	-16.2	-7.73	2.0	-15.1	-18.3	-4.8	-1.3
1971	-26.2	-27.1	-27.8	-19.39	6.3	-26.1	-30.6	-11.4	-4.1
1981	-35.9	-41.1	-37.9	-31.27	16.7	-35.8	-41.2	-17.7	-6.1

Figure 2.7.2.4: Model for India compared to the spline



x - fitted point

Other emerging nations have been examined and the results of fitting are portrayed in summary form in Table 2.7.2.3.

Table 2.7.2.3: Summary of Models of the Populations of Emerging Nations

Population	Model	Parameter	Standard Errors	Variance of residuals
Algeria 1886-1931 9 counts	3 Parameter Logistic t*= yr-1881	N_∞ = 7.769 N_0 = 3.456 c = 0.035	0.769 0.106 0.008	5.19×10^{-4}
Algeria 1936-1980 10 counts	Modified Exponential t*= yr-1881	A = 5.747 B = 0.109 C = 0.048	0.479 0.053 0.0046	6.24×10^{-4}
Argentina 1869-1970 7 counts	3 Parameter Logistic t*= yr-1869	N_∞ = 37.37 N_0 = 1.74 c = 0.0357	4.01 0.08 0.0021	2.59×10^{-3}
Chile 1835-1940 11 counts	3 Parameter Logistic t*= yr-1835	N_∞ = 8.294 N_0 = 0.501 c = 0.0220	0.016 0.040 0.0020	1.92×10^{-3}
Chile 1952-1982 4 counts	3 Parameter Logistic t*= yr-1835	N_∞ = 21.228 N_0 = 0.041 c = 0.035	10.117 0.064 0.012	5.3×10^{-4}
Colombia 1770-1973 16 counts	Modified Exponential t*= yr-1770	A = 0.648 B = 0.210 C = 0.022	0.09 0.044 0.0012	1.23×10^{-2}
Egypt 1907-1983 10 counts	4 Parameter Logistic t*= yr-1900	N_∞ = 92.0 N_0 = 1.25 c = 0.045 d = 9.79	41.64 0.46 0.0069 0.79	5.43×10^{-4}
Guyana 1841-1911 7counts	3 Parameter Logistic t*= yr-1841	N_∞ = 0.362 N_0 = 0.095 c = 0.0402	0.055 0.006 0.0078	5.04×10^{-3}
Guyana 1921-1970 5 counts	4 Parameter Logistic t*= yr-1841	N_∞ = 0.521 N_0=0.0000005 c = 0.122 d = 0.294	0.02 $<10^{-6}$ 0.0054 0.002	3.5×10^{-5}
Paraguay 1886-1980 7 counts	Geometric t*= yr-1886	A = 0.30 B = 0.024	0.04 0.002	3.18×10^{-2}
Peru 1795-1972 9 counts	Modified Exponential t*=yr-1795	A = 0.844 B = 0.3304 C = 0.0207	0.204 0.111 0.00194	1.10×10^{-2}

Table 2.7.2.3 (contd.)

Nation	Model	Parameter	Standard Errors	Variance
Uruguay 1852-1975 6 counts	3 Parameter Gompertz $t^* =$ yr-1852	$N_\infty = 4.61$ $b = -3.51$ $c = 0.0159$	0.58 0.12 0.0014	2.95×10^{-3}
Venezuela 1810-1941 12 counts	3 Parameter Logistic $t^* =$ yr-1810	$N_\infty = 5.402$ $N_0 = 0.643$ $c = 0.0197$	2.388 0.072 0.005	2.18×10^{-2}
Venezuela 1950-1981 13 counts	3 Parameter Logistic $t^* =$ yr-1810	$N_\infty = 50.783$ $N_0 = 0.017$ $c = 0.0414$	15.857 0.007 0.003	1.04×10^{-4}

The sigmoids do not appear as particularly good models. However, all populations seem to be "best" described by a model containing some exponential component; this may develop into a sigmoid model. Several populations exhibit the need for escalated modelling. This suggests that there has been renewed growth over recent years resulting in no inflections having been encountered for the renewed growth which, in turn, results in the fact that the upper asymptote cannot be specified until more information is available.

The populations of Argentina, Chile, Guyana, Uruguay and Venezuela are exceptions to this, all being described adequately by sigmoids.

2.7.3 Underdeveloped Nations

These nations represent the bulk of Africa, and parts of South America and South-East Asia. The populations are characterised by being primarily rural, often nomadic and having high fertility and increasing life expectancy (see Lesthaeghe (1986), Caldwell (1982) and Gwatkin and Brandel (1982)) although there are many nations, especially in Africa, where life expectancy at birth has not increased. The economy of such nations is based essentially on families and there is little industrialisation or national care systems. The potential these nations display for rapid population growth causes much concern, particularly in view of their limited resources. This rapid population growth in recent times has come about to a large extent as a result of modernisation overcoming many traditional checks on population growth, notably widespread infertility.

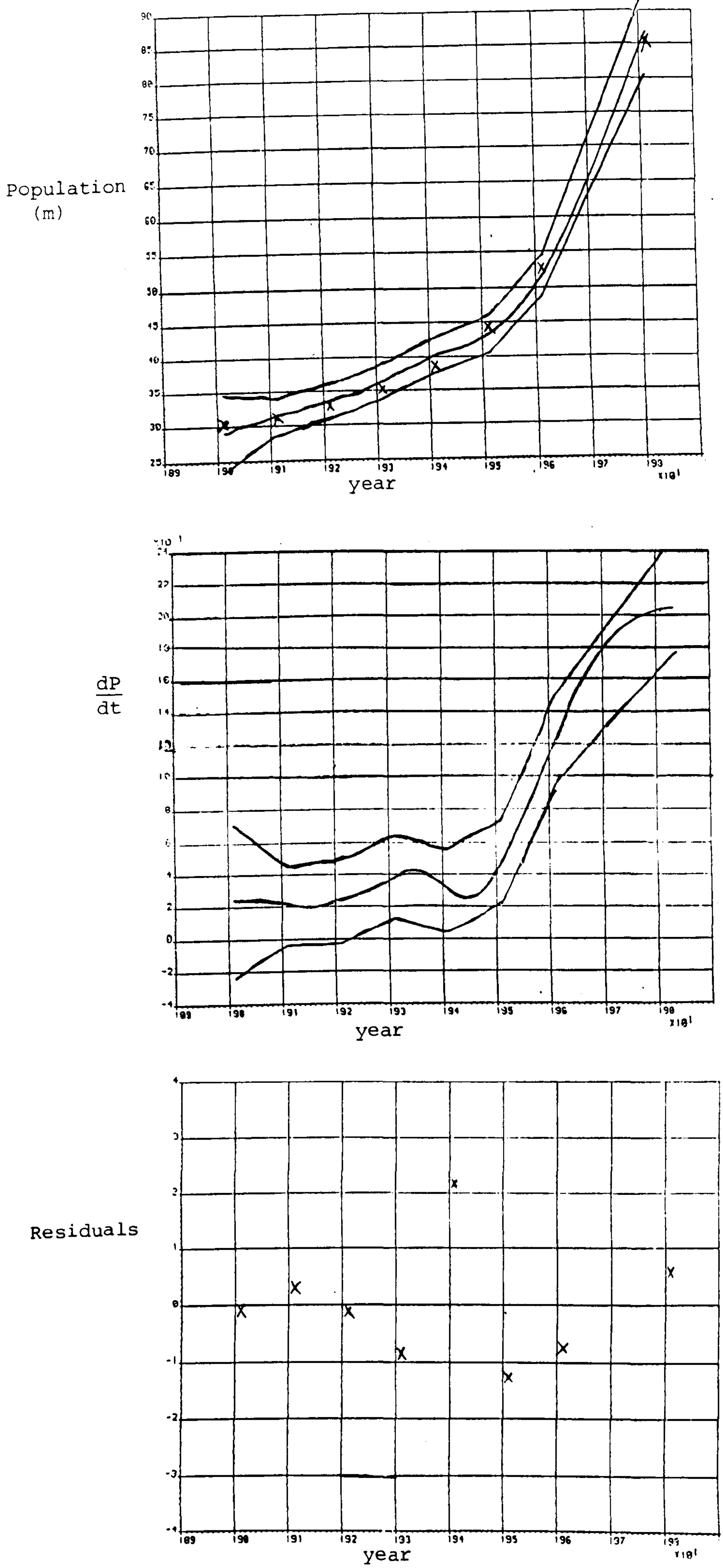
To aid policy making towards such nations good models of population growth are required. However, modelling has been hindered by the lack of reliable statistics on population growth. Using information available from census counts and United Nations estimates of total population size, the suitability of the simple trend models as descriptors of these populations will be considered, giving particular attention to the populations of Bangladesh, Kenya and Ethiopia.

2.7.3a Bangladesh 1901 to 1981

Nine census counts are available of the population of Bangladesh which is one of the economially poorest and fastest growing populations in the world.

The 95% confidence intervals on the spline are wide (see Figure 2.7.3.1), indicating that very close fits cannot be

Figure 2.7.3.1: Spline for Bangladesh



expected. It would appear from the derivative that the population of Bangladesh could have encountered a point of inflection around 1980 and growth slows from its previous geometric path.

Only the growth models were considered for this population and window analysis was not carried out as there were insufficient data to make it worthwhile. The models were fitted to the period 1901 to 1961 (7 counts) and to the extended period 1901 to 1981. These results are summarised in Tables 2.7.3.1 and 2.7.3.3. The error in the prediction of the 1981 population count are given in Table 2.7.3.2.

Table 2.7.3.2: Predictions of 1981 population count by models fitted 1901-1961

Model	% error of prediction
3 Parameter Logistic	-33.6
4 Parameter Logistic	-27.7
3 Parameter Gompertz	-38.1
4 Parameter Gompertz	-28.4
Modified Exponential	-0.25
Geometric	-33.2

The poor performance by all models, except the modified exponential, in predicting the 1981 count and the large rise in RSS from the fits to the longer period suggests that escalation of the population of Bangladesh has occurred after 1961 - perhaps this is connected to the Civil War which raged in Bangladesh in the early 1970s (this was not suggested by consideration of the spline). Alternatively the models are perhaps unsuitable for this population and alternatives need to be examined. This will be done in Section 2.8.

However, it appears that the modified exponential model derived from the 1901 to 1961 fit provides a good model and, when fitted over the period 1901 to 1981, compares favourably to the

Table 2.7.3.1: Models of the Population of Bangladesh 1901-1961

(7 counts) t* = year - 1901

Model	Parameter	Standard errors	Variance of Residuals	Sum of squares of Residuals RSS	Indexed Costed RSS
3 Parameter Logistic	N_∞ = 3918 N_0 = 28.395 c = 0.00888	48117487 1.162 0.0118	2.26×10^{-3}	16.36	2.66
4 Parameter Logistic	N_∞ = 9026 N_0 = 7.52 c = 0.022 d = 21.68	Failed to converge	2.17×10^{-3}	11.43	2.79
3 Parameter Gompertz	N_∞ = 6444 b = -5.43 c = 0.0017	237098 36.77 0.012	2.44×10^{-3}	17.75	2.88
4 Parameter Gompertz	N_∞ = 4033 b = -6.64 c = 0.00455 d = 23.94	Failed to converge	2.21×10^{-3}	11.65	2.84
Modified Exponential	A = 21.658 B = 7.55 C = 0.022	6.93 6.22 0.0103	1.63×10^{-3}	11.42	1.86
Geometric	A = 28.40 B = 0.00886	16.06 0.0008	1.81×10^{-3}	16.35	1.33
Linear	A = 28.02 B = 0.318	1.10 0.0363	2.71×10^{-3}	25.06	2.04
Quadratic	A = 29.23 B = 0.137 C = 0.0033	1.10 0.098 0.0017	1.74×10^{-3}	12.31	1.0
Cubic	A = 28.96 B = 0.253 C = -0.0024 D = 0.00067	3.68 0.052 0.0010 0.00008	2.09×10^{-3}	11.00	2.68

Table 2.7.3.3: Models of the Population of Bangladesh 1901-1981 (9 counts)

$t^* = \text{year} - 1901$

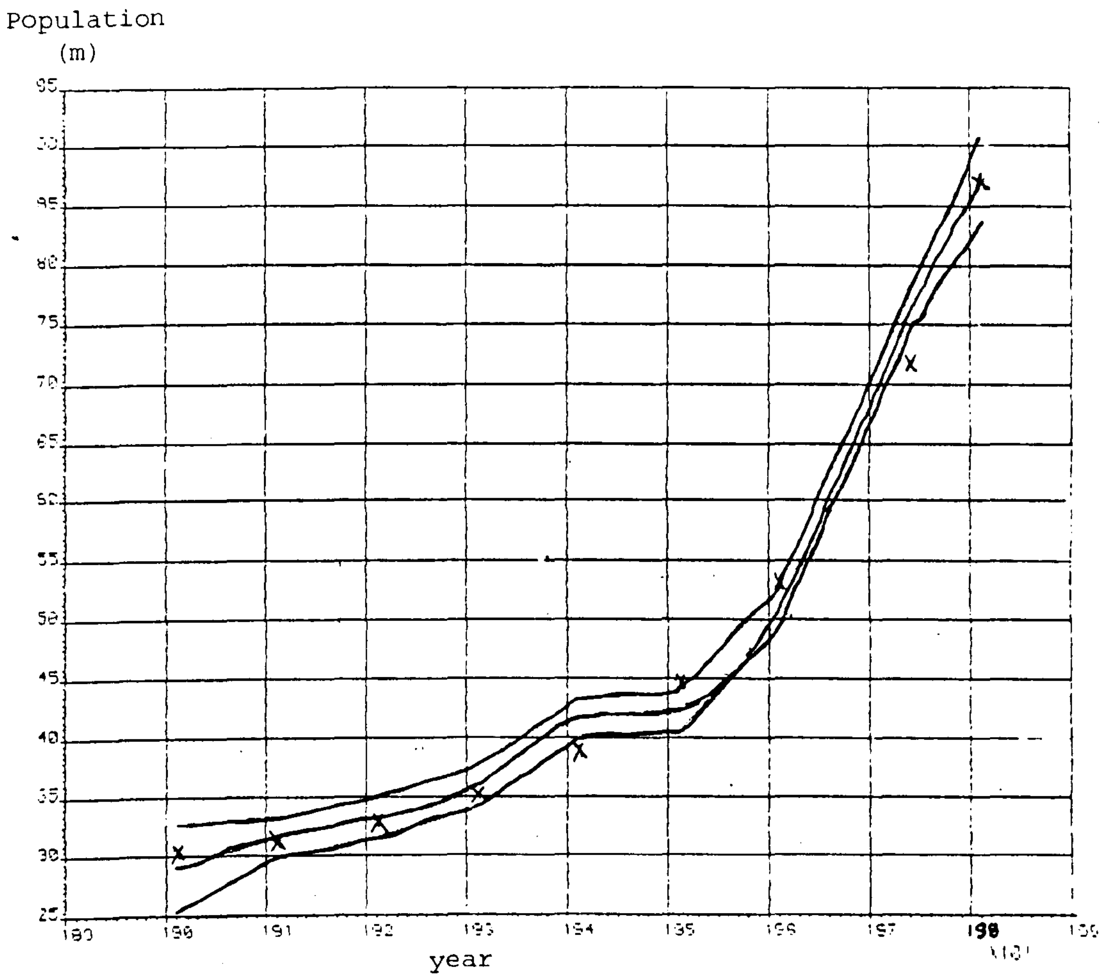
Model	Parameter	Standard errors	Variance of Residuals	Sum of squares of Residuals RSS	Indexed Costed RSS
3 Parameter Logistic	N_∞ = 124096 N_0 = 25.76 c = 0.0130	308562184 2.71 0.0118	1.78×10^{-2}	411.910	9.81
4 Parameter Logistic	N_∞ = 24444 N_0 = 1.82 c = 0.0437 d = 28.33	Failed to converge	3.39×10^{-3}	41.971	1.50
3 Parameter Gompertz	N_∞ = 22363 b = -6.77 c = 0.00203	855157 38.17 0.0125	2.00×10^{-2}	488.686	11.64
4 Parameter Gompertz	N_∞ = 15237 b = -9.95 c = 0.0072 d = 29.69	Failed to converge	4.01×10^{-3}	44.657	1.60
Modified Exponential	A = 28.322 B = 1.822 C = 0.0436	1.759 0.819 0.0058	2.82×10^{-3}	41.985	1.0
Geometric	A = 25.76 B = 0.0130	1.95 0.0016	1.53×10^{-2}	411.732	4.90
Linear	A = 25.37 B = 0.478	3.07 0.09	2.67×10^{-2}	912.092	10.86
Quadratic	A = 30.74 B = -0.163 C = 0.0098	2.34 0.165 0.0023	8.10×10^{-3}	144.429	3.44
Cubic	A = 28.61 B = 0.460 C = -0.0142 D = 0.00022	3.71 0.095 0.0062 0.00017	2.73×10^{-3}	34.991	1.25

spline benchmark as almost all estimated points lie within the 95% confidence intervals of the spline, as can be observed from Figure 2.7.3.2. Thus the model used for Bangladesh is:

$$N_t = 28.31 + 1.878e^{0.043t*}$$

$$t* = \text{year} - 1901.0$$

Figure 2.7.3.2: Model for Bangladesh compared to the spline



x - fitted value

2.7.3b <u>Kenya</u>

Population counts of the Kenyan population are only available (reliably) from 1948. This gives nine enumerations up to 1983.

The spline and its derivative (Figure 2.7.3.3) indicate that this population is growing rapidly and has not experienced an inflection in growth rates. This suggests that there is insufficient information to allow discrimination between a sigmoid and a simple geometric model.

The results of fitting over the period up to 1971 and the entire period confirm the difficulty of discrimination. A comparison of predictions made from the earlier models with the 1976, 1980 and 1983 observations is displayed in Table 2.7.3.4.

<u>Table 2.7.3.4</u>: <u>A comparison of predictions made by the models</u>

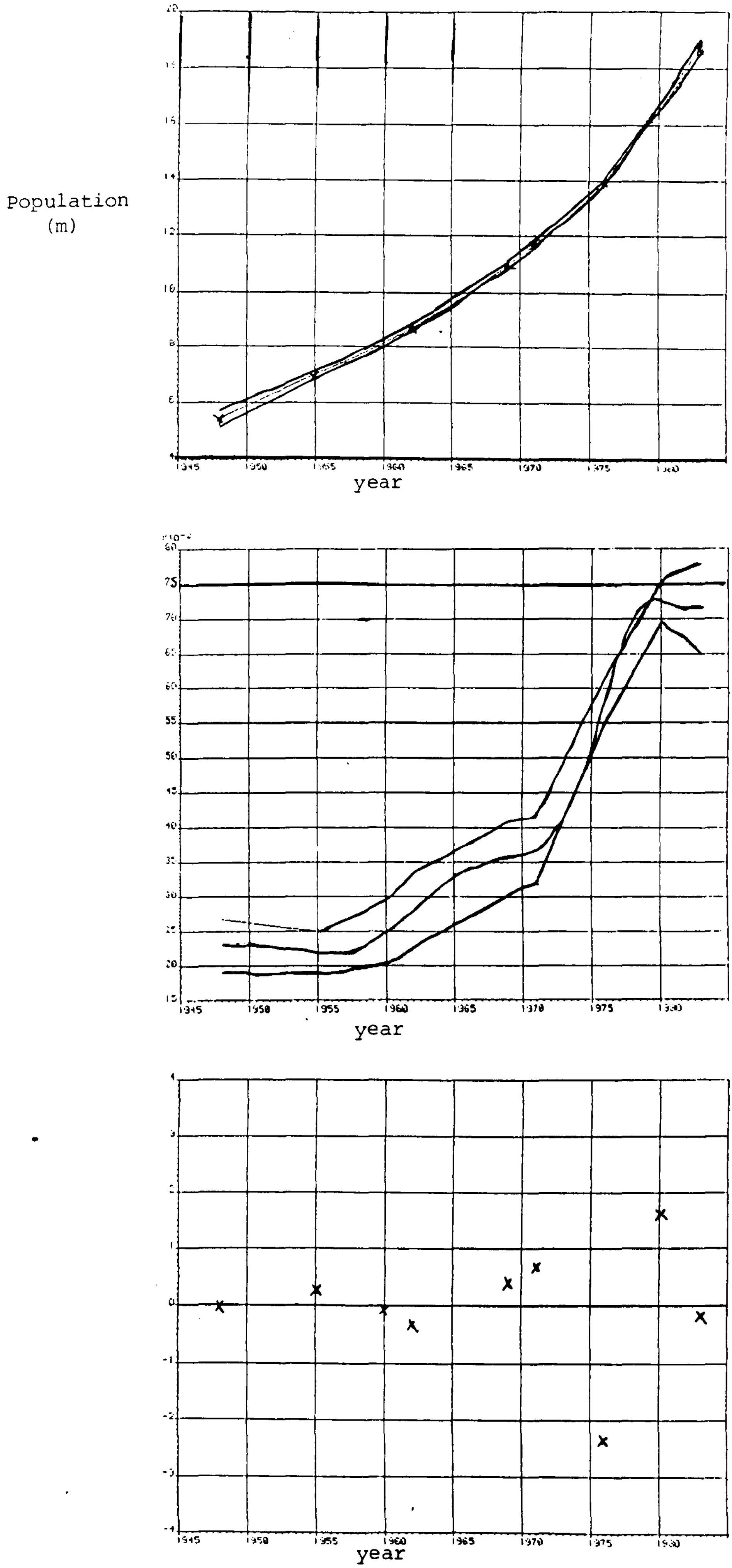
Model	Percentage error		
	1976	1980	1983
3 Parameter Logistic	-1.7	-7.7	-10.4
4 Parameter Logistic	-3.1	-8.9	-11.4
3 Parameter Gompertz	-1.6	-7.8	-10.7
Modified Exponential	-2.3	-8.0	-10.5
Geometric	-1.0	-6.1	-8.0
Linear	-9.6	-18.6	-23.7
Quadratic	-2.4	-9.0	-12.3
Cubic	2.5	0.6	1.6

These results show, as anticipated, that the sigmoids are inferior to the simpler exponential models. The geometric fits and predicts especially well. However, all models are more inferior at fitting and predicting when compared to the cubic.

Thus, until more information is available, the cubic or perhaps the geometric, should be used for describing the population of Kenya.

The model for 1948 to 1983 then is:

Figure 2.7.3.3: The spline for Kenya



$$N_t = \underset{(0.04)}{3.07} + \underset{(0.068)}{0.354t^*} - \underset{(0.0006)}{0.0097t^{*2}} + \underset{(2 \times 10^{-5})}{0.00023t^{*3}} \qquad \text{var} = 8 \times 10^{-5}$$

or

$$N_t = \underset{(0.603)}{1.615} + \underset{(0.456)}{2.772}e^{\underset{(0.003)}{0.042}t^*} \qquad \text{var} = 3.5 \times 10^{-4} \qquad t^* = \text{yr} - 1940$$

2.7.3c Ethiopia

Ethiopia is one of the least developed nations in the world and only a few semi-reliable enumerations of the population are available for this mainly rural semi-nomadic population. Eight counts covering the period 1955 to 1983 have been used.

Examination of the growth rates and graph of the Ethiopian population give no indication of inflection or even approaching inflection which suggests that the geometric model will be the most suitable.

The models were fitted up to 1977 and predictions compared with the 1980 and 1983 observed population in Table 2.7.3.5.

Table 2.7.3.5: Comparison of Predictions and Observations

Model	Percentage error	
	1980	1983
3 Parameter Logistic	-3.05	-5.43
4 Parameter Logistic	-0.35	-0.95
3 Parameter Gompertz	-3.52	-6.18
Modified Exponential	-0.35	-0.94
Geometric	-3.05	-5.42
Linear	-5.35	-8.90
Quadratic	-1.14	-2.59
Cubic	0.69	1.21

The experience of attempting to model Ehtiopia is rather similar to that of Kenya in that a sensible parameterisation of sigmoids cannot be obtained. The cubic is good but the best model of the population of Ethiopia is the modified exponential:

$$N_t = 13.07 + 4.89e^{0.044t^*}, \qquad t^* = \text{year} - 1950$$

Table 2.7.3.6 presents a summary of fits of the simple models to other underdeveloped nations.

Table 2.7.3.6: "Best" Models for Underdeveloped Nations

Nation	Model	Parameter	Standard Errors	Variance
Chad 1936-1980 5 counts	3 Parameter Gompertz t*= yr-1900	N_∞ = 7.677 N_0 = -4.06 c = 0.0247	2.47 0.46 0.0078	1.84×10^{-3}
Honduras 1901-1980 13 counts	Modified Exponential t*= yr-1901	A = 0.333 B = 0.164 C = 0.037	0.055 0.039 0.0032	4.3×10^{-3}
Laos 1921-1970 6 counts	4 Parameter Logistic t*= yr-1901	N_∞ = 2.58 N_0 = 0.0009 c = 0.139 d = 0.84	0.60 0.0017 0.0367 0.06	4.18×10^{-3}
Sierra Leone 1903-1983 10 counts	Geometric t*= yr-1901 Evidence of escalation	A = 1.115 B = 0.013	0.062 0.001	8.12×10^{-3}
Sierra Leone 1901-1956 6 counts	3 Parameter Logistic t*= yr-1901	N_∞ = 2.18 N_0 = 1.05 c = 0.0531	0.12 0.38 0.0098	1.48×10^{-3}
Sierra Leone 1963-1983 4 counts	Geometric t*= yr-1901	A = 0.495 B = 0.024	0.056 0.015	5.4×10^{-4}
Tanganyika 1921-1980 11 counts	Modified Exponential t*= yr-1900	A = 2.97 B = 0.54 C = 0.0416	0.32 0.13 0.003	1.32×10^{-3}
Tanganyika 1921-1960 7 counts	Geometric t*= yr-1900	A = 2.71 B = 0.0201	0.09 0.0008	7.91×10^{-4}
Tanganyika 1967-1980 5 counts	Geometric t*= yr-1900	A = 1.25 B = 0.0334	0.076 0.0009	1.93×10^{-4}
Uganda 1911-1979 9 counts	Modified Exponential t*= yr-1900	A = 2.344 B = 0.239 C = 0.0475	0.125 0.05 0.0026	1.02×10^{-3}

Reviewing this Table illustrates that the simple exponential models are as good as any for poor high fertility nations.

Conclusion

For poor high fertility nations the simple geometric or modified exponential models give the "best" descriptions of the historical growth of population in these countries.

However, an observation is made that there is great similarity between the geometric and the three parameter logistic and between the modified exponential (with positive parameters) and the four parameter logistic models. It is found that the models are similar in relation to the magnitude of parameters (excepting N_∞ for the logistics), and they fitted with equal precision.

A similar observation is also made when emerging nations are examined. This arises due to the population not encountering a point of inflection and so the upper asymptote (N_∞) cannot be specified and so the logistic "behaves" as an exponential model.

Thus the inferior fits of the logistics cannot be taken as evidence to refute them in the long run. If the logistic models are to be used for prediction, then one might wish to use judgment to set by other means an upper asymptote based on the population size when (or if) the growth of population starts to slow.

2.7.4 Island Populations

Table 2.7.4.1 displays summaries of the "best" fitting simple trend models to Island populations.

Out of the 22 Island populations examined, 18 were found to be best fitted by sigmoids, although seven required escalation, the remaining being "best" described by the modified exponential which may in the future follow a four parameter logistic path.

Island populations, with strict limits to land area, thus do appear to grow in a sigmoid manner unless some dramatic external influences are imposed such as major enforced emigration or catastrophic natural disaster.

Table 2.7.4.1: Summary of Best Models of Island Populations

Nation	Model	Parameter	Standard Errors	Variance
Bahamas 1851-1970 13 counts	Modified Exponential t*= yr-1851	A = 0.0313 B = 0.00287 C = 0.0307	0.005 0.0023 0.0072	2.26×10^{-2}
Barbados 1851-1901 6 counts	3 Parameter Logistic t*= yr-1851	N_∞ = 0.258 N_0 = 0.137 c = 0.020	0.054 0.002 0.008	3.65×10^{-4}
Barbados 1921-1980 5 counts	3 Parameter Logistic t*= yr-1851	N_∞ = 0.329 N_0 = 0.055 c = 0.0216	0.115 0.031 0.013	1.56×10^{-3}
Bermuda 1843-1980 15 counts	Modified Exponential t*= yr-1840	A = 8.15 B = 2.05 C = 0.0239	0.65 0.37 0.0015	2.57×10^{-3}
Cook Islands 1902-1971 13 counts	4 Parameter Gompertz (popn $\times 10^3$) t*= yr-1900	N_∞ = 26.21 b = -5.32 c = 0.0279 d = 8.10	5.31 0.47 0.0044 0.19	3.00×10^{-4}
Cuba 1774-1887 8 counts	4 Parameter Gompertz t*= yr-1770	N_∞ = 1.71 N_0 = 0.072 c = 0.046 d = 0.086	0.09 0.015 0.0036 0.019	8.2×10^{-4}
Cuba 1919-1970 5 counts	Geometric t*= yr-1770	A = 0.134 B = 0.0207	0.020 0.0009	1.12×10^{-3}
Cyprus 1881-1973 9 counts	4 Parameter Logistic t*= yr-1881	N_∞ = 0.969 N_0 = 0.064 c = 0.0304 d = 0.124	0.438 0.031 0.0090 0.034	8.26×10^{-4}
Fiji 1881-1976 10 counts	4 Parameter Logistic t*= yr-1881	N_∞ = 0.794 N_0 = 0.0027 c = 0.0647 d = 0.118	0.195 0.0014 0.008 0.004	1.58×10^{-3}

Table 2.7.4.1 (contd.)

Nation	Model	Parameter	Standard Errors	Variance
French Polynesia 1926-1971 8 counts	Modified Exponential t*= yr-1920	A = 0.0218 B = 0.0112 C = 0.042	0.0025 0.002 0.003	4.0×10^{-3}
Grenada 1851-1970 11 counts	3 Parameter Logistic t*= yr-1851	N_∞ = 0.112 N_0 = 0.0298 c = 0.0206	0.022 0.002 0.005	6.2×10^{-3}
Gilbert & Ellis Islands 7 counts	Modified Exponential t*= yr-1900	A = 0.0288 B = 0.00087 C = 0.049	0.0018 0.0007 0.011	1.5×10^{-3}
Guam 1901-1980 9 counts	4 Parameter Logistic t*= yr-1900	N_∞ = 0.102 N_0 = 0.0004 c = 0.0975 d = 0.0104	0.026 0.00048 0.0269 0.0017	2.65×10^{-2}
Jamaica 1844-1921 8 counts	4 Parameter Logistic t*= yr-1840	N_∞ = 0.657 N_0 = 0.046 c = 0.052 d = 0.322	0.102 0.024 0.011 0.032	5×10^{-4}
Jamaica 1921-1980 6 counts	4 Parameter Logistic t*= yr-1840	N_∞ = 3.176 N_0 = 0.075 c = 0.0293 d = 0.204	1.65 0.14 0.015 0.44	1.2×10^{-4}
Madagascar 1900-1940 5 counts	3 Parameter Logistic t*= yr-1900	N_∞ = 4.17 N_0 = 2.27 c = 0.07	0.325 0.104 0.022	2.16×10^{-3}
Madagascar 1950-1970 4 counts	3 Parameter Logistic t*= yr-1900	N_∞ = 61.72 N_0 = 1.25 c = 0.0256	18.65 0.03 0.0008	10^{-6}
Philippines 1903-1980 9 counts	Modified Exponential t*= yr-1901	A = 4.16 B = 3.36 C = 0.033	0.51 0.37 0.0015	6.80×10^{-4}
Samoa (American) 1900-1970 8 counts	4 Parameter Logistic (popn × 10^3) t*= yr-1900	N_∞ = 41.29 N_0 = 1.72 c = 0.0471 d = 4.06	32.85 1.51 0.0218 1.77	2.58×10^{-3}

Table 2.7.4.1 (contd.)

Nation	Model	Parameter	Standard Errors	Variance
Samoa (Western) 1900-1971 12 counts	4 Parameter Gompertz t*= yr-1900	N_∞ = 0.27 b = -8.82 c = 0.0324 d = 0.03	0.11 1.69 0.008 0.0012	2.5×10^{-3}
Seychelles 1891-1931 5 counts	3 Parameter Logistic t*= yr-1891	N_∞ = 31.35 N_0 = 15.90 c = 0.0454	2.68 0.33 0.0094	4.53×10^{-4}
Seychelles 1947-1977 4 counts	Modified Exponential t*= yr-1891	A = 29.36 B = 0.198 C = 0.0595	2.07 0.16 0.009	2.26×10^{-4}
Sri Lanka 1850-1921 8 counts	3 Parameter Logistic t*= yr-1850	N_∞ = 7.64 N_0 = 1.58 c = 0.024	1.66 0.04 0.0033	9×10^{-4}
Sri Lanka 1931-1981 6 counts	4 Parameter Logistic t*=yr-1850	N_∞ = 12.16 N_0 = 0.0003 c = 0.092 d = 4.73	0.59 0.0002 0.0057 0.12	10^{-4}
St. Lucia 1851-1901 7 counts	3 Parameter Logistic t*= yr-1851	N_∞ = 0.262 N_0 = 0.024 c = 0.017	0.041 0.007 0.004	1.1×10^{-3}
St. Lucia 1921-1970 4 counts	4 Parameter Logistic t*= yr-1851	N_∞ = 0.123 N_0 = 0.0003 c = 0.0484 d = 0.043	0.037 0.0001 0.0067 0.004	10^{-4}
St. Vincent 1851-1970 12 counts	Modified Exponential t*= yr-1851	A = 0.0288 B = 0.0038 C = 0.023	0.0042 0.0027 0.006	8.11×10^{-3}
Trinidad and Tobago 1851-1980 13 counts	3 Parameter Logistic t*= yr-1851	N_∞ = 2.78 N_0 = 0.084 c = 0.0231	1.10 0.004 0.002	5.22×10^{-3}
Tonga 1901-1976 8 counts	4 Parameter Logistic (popn $\times 10^3$) t*= yr-1901	N_∞ = 95.02 N_0 = 0.91 c = 0.076 d = 20.61	10.25 0.27 0.0066 0.62	6.06×10^{-4}

2.7.5 Quebec 1851-1981

Over the period there have been fourteen decennial census counts. This population is of special interest as Quebec has been a relatively self-contained population and has experienced little migration.

The spline, its derivatives and residuals support the use of sigmoid models (see Figure 2.7.5.1).

However, the cusum plot of this population, illustrated in Figure 2.7.5.2, does not show the inverted U-shape that is characteristc of a logistic type process. The problems of using a three parameter logistic are confirmed when the window analysis is considered. From the window analysis good fits for the three parameter logistic are found between 1851 and 1901, and between 1901 and 1981. These have upper asymptotes of around 4 million and 23.5 million respectively. The latter asymptote is unexpectedly high. For the four parameter logistic model no requirement for escalation is detected. However, a conclusion drawn from analysis of the backward moving window is that a significantly closer fit is obtained for the four parameter logistic if the 1851 census point is omitted. This gives an upper asymptote of 10.45 million.

After comparing different trend models the most suitable model for Quebec (1861 to 1981) is the four parameter logistic model parameterised as:

Parameter	Standard Error	
$N_\infty = 9.51$	1.49	
$N_0 = 0.13$	0.04	
$\alpha = 0.036$	0.003	$t^* = \text{year} - 1851$
$d = 0.94$	0.23	
variance $= 7.5 \times 10^{-4}$		

Figure 2.7.5.1: The spline for Quebec

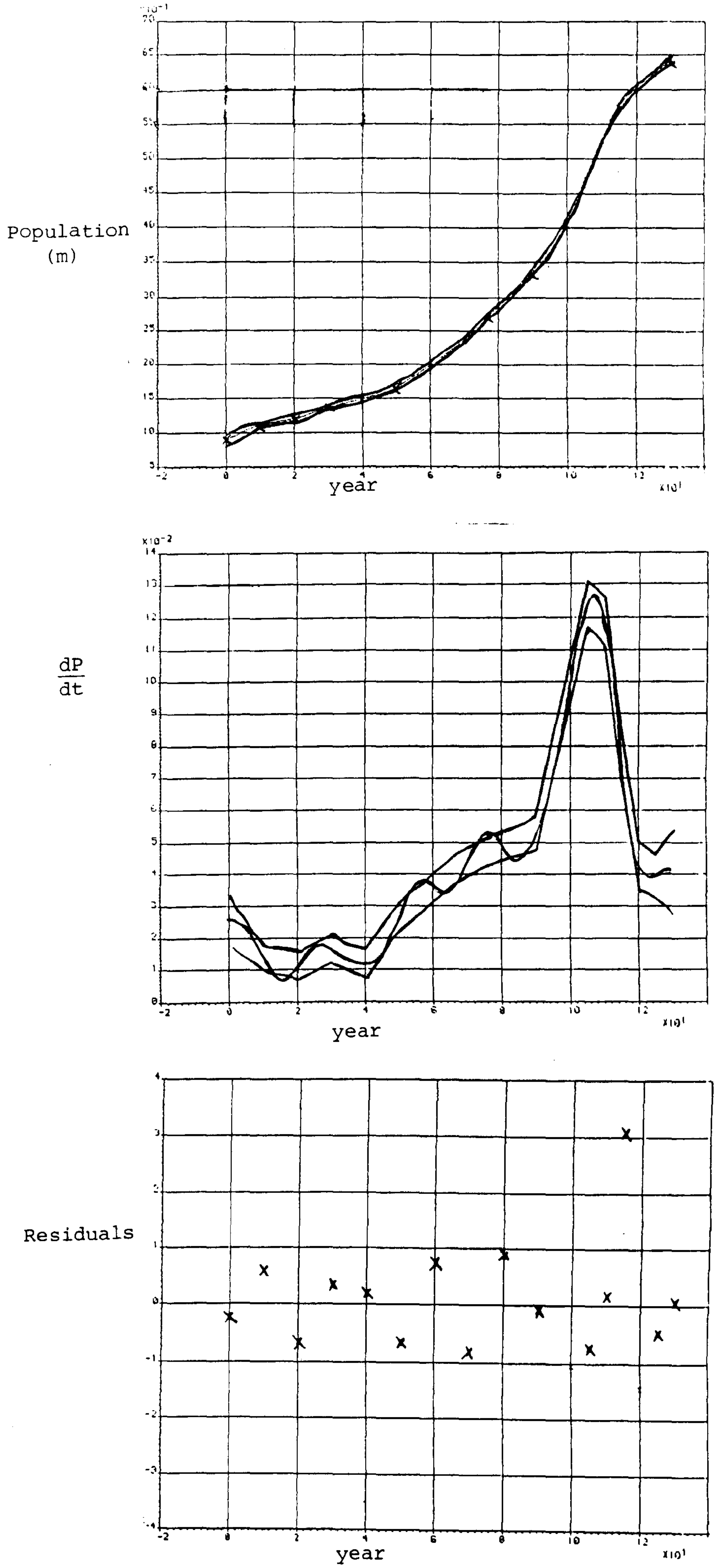
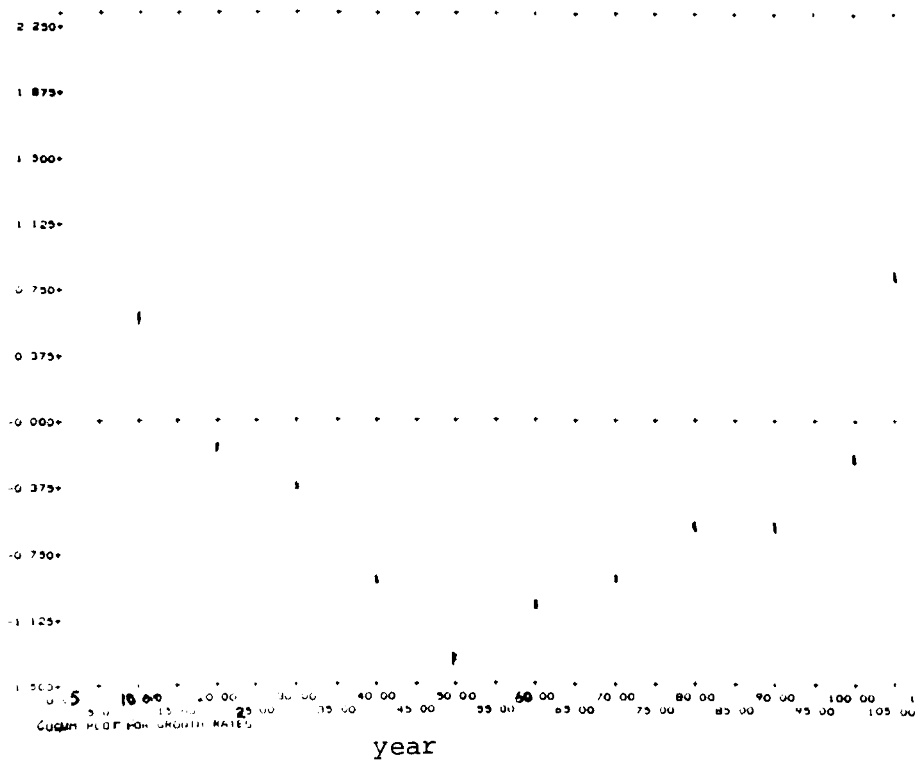


Figure 2.7.5.2: Cusum of Quebec's growth rates



Thus it appears that sigmoids are good for populations where growth has come about mainly through birth minus deaths and where migration has had little influence, (ie natural increase).

2.7.6 Models of Component Populations

The population of Great Britain over the period 1801 to 1981 can successfully be modelled by a logistic model. It is of interest to investigate if the component populations of England, Wales and Scotland can be described by a similar model.

After exploratory analysis the most suitable models are found to be:

(I) England (1801 to 1981) S.Es.

- 4 Parameter logistic

		S.Es.
N_∞	= 50.51	(1.92)
N_0	= 6.06	(0.63)
c	= 0.0218	(0.0011)
d	= 2.27	(0.70)
variance	= 2.2×10^{-4}	(7.0×10^{-5})

(II) Wales (1801 to 1981)

- 4 Parameter logistic

N_∞	= 2.59	(0.21)
N_0	= 0.23	(0.08)
c	= 0.0293	(0.0040)
d	= 0.38	(0.10)
variance	= 2.43×10^{-3}	(8×10^{-4})

There are indications that an improved model is:

$$N_t = \begin{cases} \dfrac{4.28}{1 + \left(\dfrac{4.28}{0.59} - 1\right)e^{-0.0166t^*}} & \text{for 1801 to 1901} \\[2ex] \dfrac{3.52}{1 + \left(\dfrac{3.52}{1.92} - 1\right)e^{-0.006t^*}} & \text{for 1931 to 1981} \end{cases}$$

The de-escalated model reflects the depressed growth rates of the Welsh population after 1910 as a result of high emigration, mainly to England, which appeared beneficial as a consequence of declining industrial and economic prosperity.

(III) England and Wales S.Es.

- 4 Parameter logistic

		S.Es.
N_∞	= 52.77	(1.96)
N_0	= 6.19	(0.65)
c	= 0.024	(0.001)
d	= 2.75	(0.72)
variance	= 2.1×10^{-4}	(10^{-4})

(IV) Scotland S.Es.

- 4 Parameter logistic

$$
\begin{aligned}
N_\infty &= 4.54 && (0.32)\\
N_0 &= 0.75 && (0.20)\\
c &= 0.0276 && (0.0029)\\
d &= 0.88 && (0.22)\\
\text{variance} &= 6.7 \times 10^{-4} && (2.2 \times 10^{-4})
\end{aligned}
$$

(Note in the above models, as elsewhere, the upper asymptote and initial populations are at $N_\infty + d$ and $N_0 + d$ respectively.)

This model performs well but exploratory analysis indicates that there is a definite requirement for de-escalation after 1931 when the Scottish population fell by 1.3%. The need for de-escalation was also detected by Leach (1981) and is a consequence of a relatively poorer economic situation which has encouraged high emigration.

The optimal model for Scotland then is:

$$
N_t = \begin{cases} \dfrac{6.80}{1 + \left(\dfrac{6.80}{1.61} - 1\right)e^{-0.017t^*}} & \text{for 1801 to 1931} \\[2ex] \dfrac{5.28}{1 + \left(\dfrac{5.28}{0.65} - 1\right)e^{-0.035t^*}} & \text{for 1951 to 1981 (obtained from fits to yearly estimates)} \end{cases}
$$

$$t^* = \text{year} - 1801$$

(Both the above parts of the model are in reasonable agreement for the 1941 population size at around 5m.)

The models of the separate populations are combined and are found to compare closely with the model of total population. The results are presented in Table 2.7.6.1.

Table 2.7.6.1: Models Combined

Models

	III	I + II	Great Britain	III + IV	I + II + IV
$t^* = 0$	8.94	8.94	10.56	10.57	10.57
$t = \infty$	55.52	55.75	60.68	60.94	61.17
d	2.75	2.65	3.32	3.63	3.53

The Roman numerals in Table 2.7.6.1 refer to the population models listed above.

Thus it would seem that component populations can be modelled separately and combined to give a model of total population. This confirms Reed and Pearl's (1927) findings that combining models of component populations approximates models of the total population (Reed and Pearl investigated the white and Negro component populations in the USA.)

In the component populations of Great Britain migration has a considerable influence as a component of population growth. To get some idea of the effect of migration the component populations are reconstructed to account for migration. The reconstruction was based on the yearly accounting formula:

$N_t = N_{t-1} + N_{t-1} \times$ (crude birth rate - crude death rate at time t)

This gave the populations of England and Wales and of Scotland as:

Year	1841	1851	1861	1871	1881	1891	1901	1911	1921
England & Wales	15.91	17.70	19.99	22.71	26.25	29.81	33.44	37.46	40.41
Scotland	-	-	3.06	3.47	3.96	4.45	4.96	5.50	5.92

Year	1931	1941	1951	1961	1971	1981
England & Wales	42.57	43.54	45.34	47.36	50.05	50.56
Scotland	6.27	6.49	6.81	7.16	7.39	7.43

Modelling these reconstructed populations gives the models:

(I) <u>England and Wales</u>

	Parameters	S.Es.
- 4 Parameter logistic	$N_\infty = 41.37$	(1.70)
	$N_0 = 5.57$	(0.09)
	$c = 0.0349$	(0.0023)
	$d = 10.24$	(1.03)
$t^* =$ year - 1841	variance $= 2.1 \times 10^{-4}$	(8×10^{-5})

(II) Scotland

- 3 Parameter logistic

$$N_\infty = 8.20 \quad (0.11)$$
$$N_0 = 3.03 \quad (0.022)$$
$$c = 0.0242 \quad (0.0007)$$
$$\text{variance} = 9 \times 10^{-5} \quad (3 \times 10^{-5})$$

Thus, as anticipated, by correcting for migration, one obtains models that give higher populations for Scotland and lower populations for England and Wales, although the saturation population of models I and II combined is similar to the model of the population of Great Britain (i.e. $N_\infty + d = 59.81$).

Thus, since combining models of component populations gives a close approximation to the model of total population, then perhaps a way of obtaining models of component populations is to take the appropriate proportion of the total population. However, due to the effect of migration between regions and the stochastic nature of population growth (see the simulations in section 2.5), this methodology is likely to be of limited value when the size of the component population is small, unless the component population is a biologically or a culturally distinct group.

In this context it should be noted that aggregating sigmoids of the same type does not, except under exceptional circumstances, add up to one of the same type. For a discussion of aggregation of sigmoids, see Reed and Pearl (1927), Yule (1925) and Stone (1979).

2.7.7 England 1541-1981

Wrigley and Scofield (1981) reconstructed the English population over the period 1541 to 1871. This series, sampled every five years, when added to the census counts, gives a considerably longer series of population counts for which the modelling can be investigated.

Figure 2.7.7.1 illustrates the spline fitted to this series, its derivative and residuals from the fit. This exploration of the series suggests that a sigmoid is suitable for modelling the English population. However, considering the derivative of the spline, a single sigmoid may not be appropriate as from 1541 to around 1700 the growth rate (as represented by the derivative) slows and stationarity is approximated. After this a shape suggesting a sigmoid is apparent with point of inflection around 1920. If a symmetrical sigmoid is suitable, then the upper asymptote may be expected to be in the region of fifty million.

Fitting over the entire period (78 counts) gives the four parameter logistic as the "best" model out of those considered. This model is characterised as:

	Parameter	S.E.
N_∞ =	47.62	0.86
N_0 =	0.012	0.003
c =	0.0235	0.0006
d =	4.18	0.10
variance =	1.53×10^{-2}	

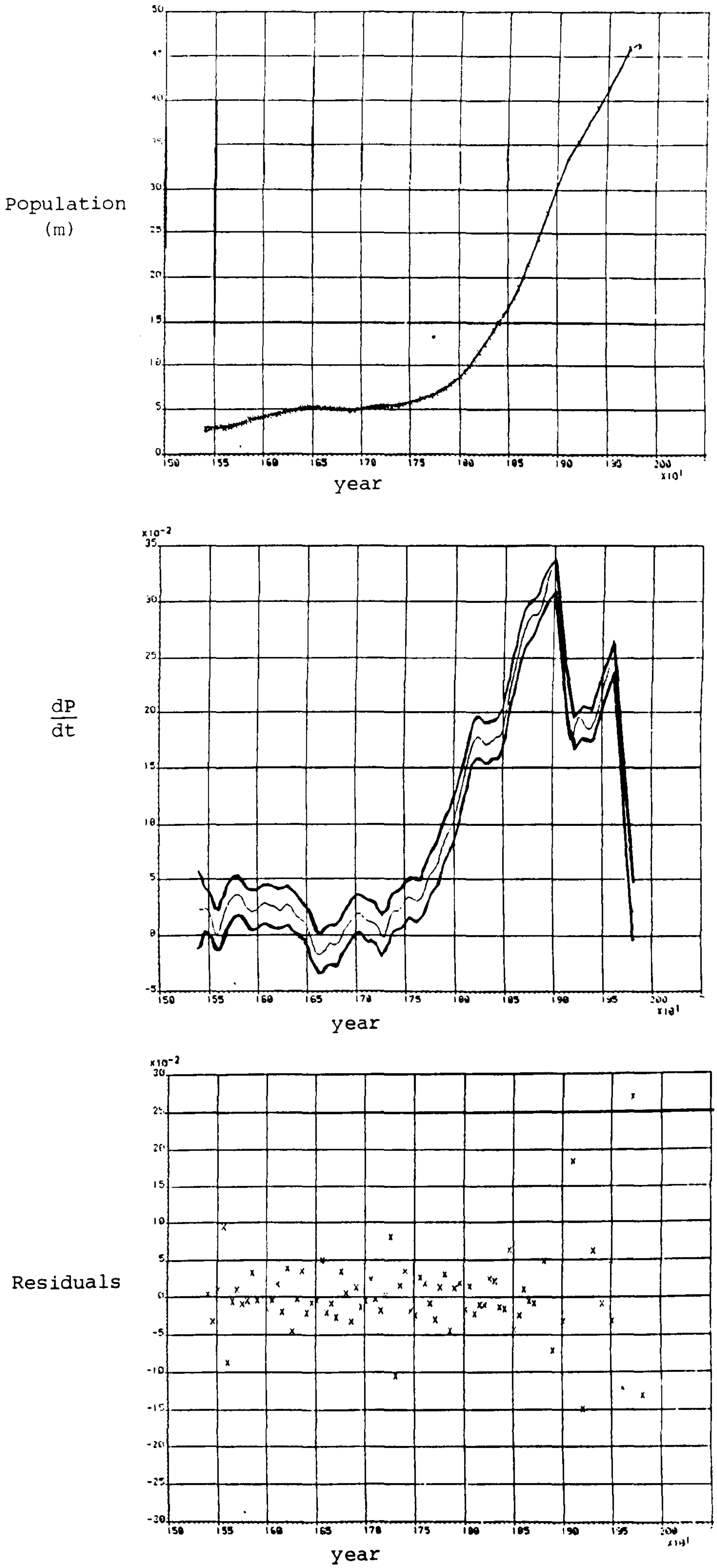
which gives a close description of the population series. This is compared to models fitted to the period 1541 to a period prior to 1700 and from 1701 to 1981, as was suggested by the spline analysis. The four parameter logistic model emerged as the most suitable from the trend models examined. These have parameters as follows:

Model for the period 1541 to 1696 (30 counts)

	Parameter	S.E.
N_∞	2.370	0.12
N_0	0.128	0.048
c	0.056	0.006
d	2.724	0.09
variance	6.3×10^{-4}	

(This model and its parameters are stable and fit well up to 1851)

Figure 2.7.7.1: The spline of England 1541-1981



Model for the period 1701 to 1981 (46   counts)

	Parameter	S.E.
N_∞	46.64	0.62
N_0	0.009	0.002
c	0.024	0.0005
d	4.41	0.14
variance	5.2×10^{-4}	

As can be seen from considering the above models, there is little difference in the parameters for the model fitted over the entire range of the data and the model fitted over the period 1701 to 1981. However, prior to 1611, the observed population is less than the lower asymptote of around 4.41 which is unsatisfactory. Hence the two models described above for the periods 1541 to 1696 and 1701 to 1981 should be used and give a satisfactory escalated logistic model.

2.8 Global Sigmoid Models

In Section 2.7 several cases were found where the simple single sigmoid models were found to be inadequate. There are other populations such as Ireland, and the Netherlands over the period where growth rates fluctuate so widely that the simple sigmoid models are clearly ruled out. It is the aim in this Section to examine whether sigmoids of a more general class can be used to overcome the inadequacy of the simpler models when "phase changes" in rate of population growth is encountered.

Stone (1979) suggests ways of modelling populations that would require escalation or de-escalation of growth curves within a continuous function. He suggests:

(a) a moving upper bound which is given by the expression:

$$N_t = \frac{W}{1 + e^{\alpha-\beta t} + ce^{a-bt}}$$

which can be derived from the differential equation:

$$\dot{N} = \beta N\left(1 - \frac{N}{\sigma_t}\right).$$

This is similar to the differential equation for the three parameter logistic, except that $\sigma_t = \frac{W}{1 + m_t^{-\theta}}$ which shows that σ_t has an upper bound W, and depends on a variable m_t, which Stone assumes to grow exponentially, i.e. $m_t = m_o e^{pt}$.

Such a model, with its many parameters, is much more flexible than the simple three or four parameter sigmoids and would appear to offer a good model for the Netherlands where there have been periods of population growth followed by a trend around zero population growth, then land reclamation and renewed population growth.

(b) a variable intrinsic rate of growth - this is given by the expression:

$$N_t = \frac{N_\infty}{1 + e^{\alpha-\psi(t)}}$$

which can be obtained from the differential equation:

$$\dot{N} = \psi(t)N(1 - N/N_\infty).$$

In this case the intrinsic rate of growth is no longer a constant but is given by a time dependent function having an integral $\psi(t)$. If $\psi(t)$ is a polynomial or a wave type function, then, depending on its order, a model of growth and decline may be obtained.

This model is appealing as it means that, in the absence of migration, crude birth rates minus crude death rates need not be a linear function, as is suggested by the simple logistic model.

This function was originally used for population modelling by Pearl and Reed (1923), and Pearl (1924) found it to be useful in modelling the populations of Germany 1816 to 1910 and Japan 1732 to 1920.

(c) several determining variables - the only determining variable considered has been time, but more could be input into the logistic equation and, depending on their sign and magnitude, this could result in irregular growth with periods of decline. The logistic equation, with X_n determining variables, becomes:

$$N = \frac{N_\infty}{1 + e^{\alpha-(\beta_1X_1+\beta_2X_2+\beta_3X_3+\ldots+\beta_nX_n)}}$$

These variables could be economic, social, psychological, biological, etc. But there are major problems in determining their:

(i) effects

(ii) interrelationships

(iii) reliability

(iv) magnitudes

(v) importance.

Thus incorporating extra variables adds greatly to the complexity of the models used and defeats a proposition of this thesis of using simple models based on readily obtainable data, and so such extensions are not considered.

The extensions outlined in 2.8(a) and 2.8(b) are used to fit some of the populations which are not described particularly well by a single sigmoid. The fitting algorithm was based on the NAG10 library routine E04FDF. The results of these fits are displayed in Table 2.8.1 which exhibits the best (in terms of lowest sum of squares of residuals) model obtained.

These more complex models do give substantial reductions in the sum of squares of the residuals. However, one must be cautious regarding the quality of the parameter estimates since they are likely to be highly correlated with each other, and are likely prone to wide specification since the sum of squares of the residuals surface is flat in the direction of the upper asymptote (see Mar Molinero, 1984 and Figure 2.5.2). Also one must question if there are sufficient data available to support such complex modelling (see Nelder, 1961). The problems of estimation are borne out by the observation that many iterations are required to achieve convergence of the fits of these models, sometimes the convergence criteria could not be satisfied.

The 'best' model of the Netherlands was found to be the moving upper asymptote as suspected, while other notable fits are a model with variable intrinsic growth rate of a polynomial form to Ireland and a model with a sine wave form of intrinsic growth to France. These fits compared to observed populations of Ireland and the Netherlands, are displayed in Figure 2.8.1. However, extrapolations from such models give, as Figure 2.8.1 illustrates, rather implausible futures for these populations.

Due to the poor predictive ability of these models and the difficulty of fitting them and obtaining 'reasonable' parameter estimates, investigation of them will be carried no further. Rather, for cases where one simple model proves unsuitable, then modelling subsections of the time series by the simple models will be adopted.

Stone (1979) points out that, by adding an extra parameter (θ) to the differential equation of the three parameter logistic model, families of curves can be generated. This is formulated as:

$$\dot{N} = \beta N\left[1 - \left(\frac{N}{N_\infty}\right)^{1/\theta}\right]$$

which, on integration, gives:

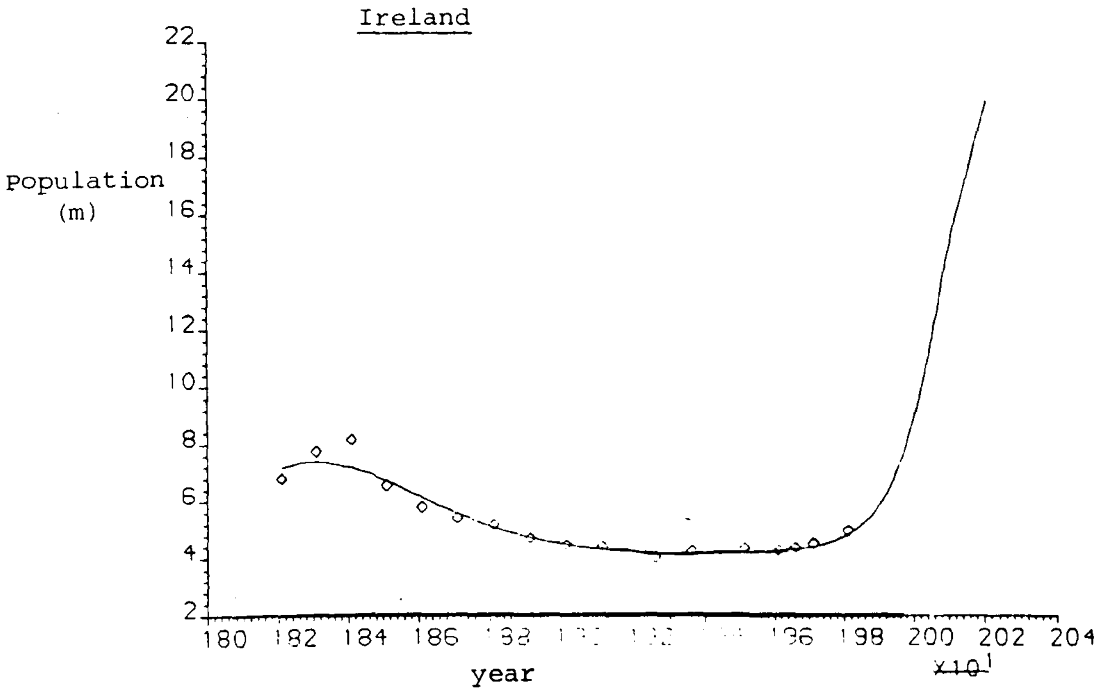
$$N = \frac{N_\infty}{\left[1 + e^{(\alpha-\beta t)/\theta}\right]^\theta}$$

The parameter, θ, determines the shape of the curve and represents the ratio of the initial growth rate to the final gap-shrinkage rate as the upper asymptote is approached.

This has a point of inflection at $\frac{N}{N_\infty} = [\theta/(1+\theta)]^\theta$ and $t = (\alpha + \theta \ln \theta)/\beta$.

The value of θ gives different sigmoids, for example, when $\theta = 1$, the model is the three parameter logistic and, when θ tends to infinity, the model is the three parameter Gompertz.

Figure 2.8.1: Models for Irish and Dutch populations



- observation

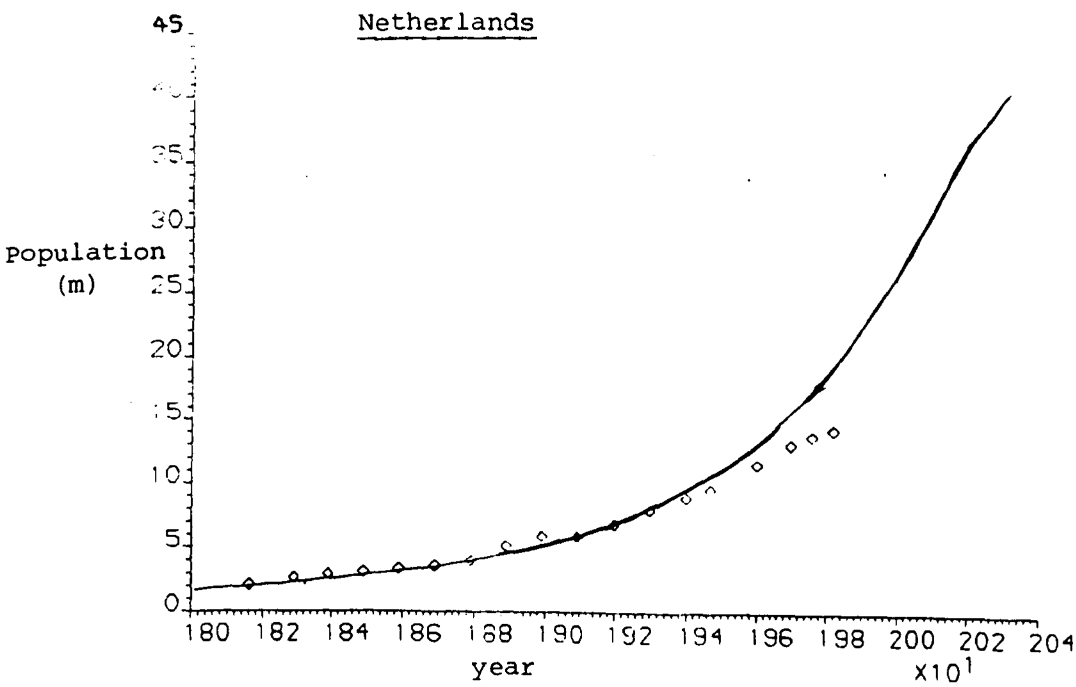


Table 2.8.1: Fits of Extended Sigmoids

Population and Model	Weighted SSR	% reduction in SSR
BRAZIL: 1808-1980; 13 counts Moving upper asymptote $N_t = \dfrac{631.27}{1+0.43e^{-0.00057t*} - 1.09e^{-6.347-0.0276t*}} - 440.532$ $t* = \text{year}-1801$	0.049	57%
CHILE: 1837-1970; 14 counts Variable intrinsic rate of growth $N_t = \dfrac{158.63}{1+83.845e^{-0.0067t*-0.00019t*^2}} - 1.38$ $t* = \text{year}-1801$	0.037	
ENGLAND: 1801-1981; 19 counts Variable intrinsic rate of growth $N_t = \dfrac{50.78}{1 + e^{1.63-0.017t*-2.56\times10^{-5}t*^2}}$ $t* = \text{year}-1801$	0.004	29%
ENGLAND AND WALES: 19 counts Variable intrinsic rate of growth $N_t = \dfrac{53.29}{1+e^{1.61-0.017t*-2.99\times10^{-5}t*^2}}$ $t* = \text{year}-1801$	0.0041	36%
FRANCE: 1801-1981; 18 counts Variable intrinsic rate of growth $N_t = \dfrac{413.2}{1+e^{176.9+178.375\sin(86.7t*+39)}}$ $t* = \text{year}-1801$	0.043	90%
INDIA: 1871-1981; 12 counts Variable intrinsic rate of growth $N_t = \dfrac{1896.21}{1+1.36e^{0.0478t*-0.00025t*^2}} + 107.40$ $t* = \text{year}-1801$	0.0045	

Table 2.8.1 (contd.)

Population and Model	Weighted SSR	% reduction in SSR
IRELAND: 1821-1981; 17 counts Variable intrinsic rate of growth $N_t = \dfrac{17.48}{1+e^{1.446-0.014t^*+8.1\times10^{-4}t^{*2}-4.18\times10^{-6}t^{*3}}} + 3.88$ $t^* = \text{year}-1821$	0.033	87%
JAPAN: 1852-1980; 12 counts Variable intrinsic rate of growth $N_t = \dfrac{184.59}{1+17.53e^{-0.0087t^*-5\times10^{-6}t^{*3}}} + 10.33$ $t^* = \text{year}-1801$	0.00078	88%
NETHERLANDS: 1816-1982; 18 counts Moving upper asymptote $N_t = \dfrac{0.225}{1+2.46e^{-0.0022t^*}-0.074e^{3.81-0.0013t^*}} - 0.238$ $t^* = \text{year}-1801$	0.030	29%
QUEBEC: 17 counts Variable intrinsic rate of growth $N_t = \dfrac{7.182}{1+e^{1.93-0.0194t^*+2.1\times10^{-4}t^{*2}-2.36\times10^{-6}t^{*3}}}$ $t^* = \text{year}-851$	0.0117	85%
SCOTLAND: 1801-1981; 19 counts Variable intrinsic rate of growth $N_t = \dfrac{5.10}{1+e^{0.80-0.25t^*+1.9\times10^{-4}t^{*2}-2.0\times10^{-6}t^{*3}}}$ $t^* = \text{year}-1801$	0.007	58%

Table 2.8.1 (contd.)

Population and Model	Weighted SSR	% reduction in SSR
U.S.A: 1790-1980; 20 counts Variable intrinsic rate of growth $N_t = \dfrac{413.0}{1+e^{4.68-0.034t^*-4.72\times 10^{5}t^{*2}}}$	0.024	48%
<u>or</u> $N_t = \dfrac{346.4}{1+e^{12.23+14.17\sin(89.85t^*-3.72)}}$ $t^* = \text{year}-1790$	0.025	46%
WALES: 1801-1981; 19 counts Variable intrinsic rate of growth $N_t = \dfrac{2.71}{1+e^{1.31-0.027t^*+2.75\times 10^{-4}t^{*2}-2.54\times 10^{-6}t^{*3}}}$ $t^* = \text{year}-1801$	0.017	69%

Models of this type, to which can be added a displacement term to give a non-zero lower asymptote, are sometimes referred to as powered logistics. Von Bertalanffy (1957) and Richards (1959) have used powered logistics to investigate animal and plant growth.

It would be useful to fit this four parameter generalisation and allow the least squares algorithm to optimise the type of model by estimating θ.

Unfortunately, fitting failed to converge. This experience was also found by Nelder (1961), McNeil (1974) and Mar Molinero (1984) and fitting was found to be of little value. To overcome this, authors have set values of θ; McNeil illustrates a methodology for estimating θ based upon an exploratory Data Analysis approach, but finds he is unable to estimate the parameters in this model to provide useful forecasts.

Nelder (1961) suggests a nested approach to fit such complex models. However, if population censuses are to be used, there is insufficient data to allow such fitting. (Nelder (1961) points out that, if five parameter models are to be fitted, then very extensive data would probably be needed to obtain reliable estimates.)

Mar Molinero (1984), commenting on this and on the problem of trying to correct the error structure for correlation, suggests that the model is overparameterised. He considers the sum of squares function to be either degenerate or too flat in the direction of θ.

Using a least squares algorithm to identify the correct model is thus unsatisfactory and is dangerous since, as Mar Molinero points out, the policy implications in terms of forecasting from the models often differ greatly.

Therefore, due to the specification problem, this general modelling is abandoned in favour of specifically setting the models.

2.9 Discussion

In this Chapter is a methodology for fitting trend curves to population data and this was adapted to fit populations for various nations from different regions.

The results can be summarised in that, regardless of economic or cultural settings, those nations which have undergone a demographic transition (i.e. a move from a high fertility/ mortality regime to a low fertility/mortality regime, see Coale (1974)) are closely described by sigmoids. Thus one could consider a sigmoid function to be a mathematical statement of the demographic transition (this point is developed further in Chapter 4 and the merit of the construct of the demographic transition is considered in Chapter 5). Many closed Island populations irrespective of economic development, are also described well by sigmoid models. After reviewing such evidence, one is left considering that perhaps Pearl (1925) was correct in referring to the logistic model as a 'law of population growth'. This view will be assessed in Chapter 5.

For emerging and less developed regions, results do not support modelling by sigmoids so well. However, for cases where sigmoids were not the best model, the geometric or the modified exponential was. These models give the same description as the logistic model. This has resulted because these populations have not undergone an inflection in their recent growth rates. Thus, given time, these populations may also follow sigmoids. (Hence, failure of a sigmoid to fit a population as well as an exponential model may be a reflection of a statistical artefact that there is insufficient data past the point of inflection, rather than an indication of theoretical inadequacy of sigmoid models.) By considering demographic theory one might postulate appropriate asymptotes for such populations by considering when (or if) growth will tend to zero. Some populations, such as the USA and France, one model was not found to be adequate as there was renewed growth. However, both periods were found to be adequately described by models fitted to each regime of growth in an escalated way, and this is preferable to fitting more complex models.

Thus it is proposed that sigmoids, particularly logistic models, are good models for interpolating and describing populations (in terms of low deviations from the fitted path). On the basis of these results one should perhaps be concerned and give further investigation to populations which do not follow a sigmoid path as they would appear to be exceptions.

Having reached the conclusion that sigmoids are good models of population, the use of such models for forecasting populations will be considered in the next Chapter.

Chapter 3

Forecasting National Populations Using Sigmoids

3.1 Introduction

Forecasts of population are required for many reasons, some of the most obvious are to aid in:-

(a) economic planning - e.g. what size will the potential workforce be 15 years from now, or how many people will be in a specific age/sex defined market?

(b) infrastructural planning - e.g. what power generating capacity will be required 30 years from now, how many houses should be built or how many health workers or teachers should be trained?

(c) military planning - this was of great concern in the past when a nation's future, to an extent, was determined by its armed might which was to a large extent governed by the population size.

Demographic issues are still of concern, particularly in West Germany where there is a worry over population decline, and in Eastern bloc nations where there is concern over changing ethnic mix of the population.

Different users of forecasts have different demands of forecasts in terms of accuracy and of lead times. For example, to plan for the commissioning of a new reservoir needs a lead time in excess of 20 years, while planning for the size of a market for consumer durables requires a lead time of only a few years. Accuracy and lead time are related in that the longer the lead time for a forecast the larger the error between forecast and observation will likely be. Keyfitz (1979) suggests that, in practice, population forecasts have large standard errors which grow exponentially as lead times increase and that forecasts for lead times in excess of 25 years are unusable. Since in the developed nations at least mortality is fairly predictable (Brass, 1974), the error in forecasting population is determined by the extent to which fertility and migration deviate from expected paths.

Official bodies, such as the Office of Population Censuses and Surveys (O.P.C.S.) in Great Britain and the US Bureau of the Census, state that they do not make forecasts but projections. However, Keyfitz (1972) criticises the practice of demographers who present a range of projections for not stating assumptions clearly and not giving potential users guidance as to how to choose appropriate projections. He points out that, when a user is compelled to "choose a likely future population there is no difference from compelling him to choose a future population from a set of random numbers". Pittenger (1978) also points out that there are an infinite number of possible projections - forecasts are a subset of these representing the most likely of the future scenarios. While acknowledging the existence of purely analytic projections, Pittenger (1978) states that "it seems that most demographic projections are functionally population forecasts, no matter the demographer's claims or caveats".

The main methodologies used in demographic forecasting have been classified by Long (1984) as:-

(i) judgmental - this is summarised well by Long and McMillan (1984) as: "in the judgmental tradition the forecaster guesses at the ultimate level of population of its growth rate, or of its rate of change. These guesses are often based on theory or experience but are not necessarily made under a formalised set of criteria". An example of a judgment based approach is the Delphi technique (see Ascher, 1979) where experts using their experience tease out likely scenarios. Judgment is part of all forecasting methods and must be recognised as an important feature.

(ii) Demographic accounting methods - this is principally the cohort-component method, where the component rates of population are disaggregated and analysed and forecasted separately. These rates are used as transition probabilities in a Markov procedure that involves updating the life table of the population (see Cox, 1970).

In these methods scenarios based on arbitrarily chosen assumptions as to how the component rates may change and judgmental methods are used to assess the most likely scenarios.

This method has the advantage of using the age structure of the population and was first applied to the English and Welsh population by Cannon in 1895 and later by Bowley (1924). The method was popularised in the USA by Welpton in 1936 and Leslie (1945, 1948) developed and presented a compact form of the method using matrices which are the basis of current cohort-component projection methods.

(iii) Explanatory Methods - theories and models of population growth are formed relating population change to other variables - often of an economic nature. From this basis forecasts can be made. Such a method is often applied to population components particularly fertility, examples are projections made by Easterlin (1978) and Butz and Ward (1979). However, in practice, such methods have been of limited use due in a large part to the lack of quantifying and difficulty of projecting explanatory variables. Also there are problems of lack of up to date information as the current cohort is incomplete.

(iv) Time Series Methods - these principally follow two lines - deterministic modelling and stochastic modelling. Deterministic modelling involves the fitting of "trend" models such as those discussed in Chapter 2. The trend function is fitted to past population counts and extrapolated into the future. Examples are the successful forecasting carried out by Pritchett (1891) who used a cubic function, Pearl and Read (1920) using a logistic function and more recently Leach (1981) also using the logistic. The earliest official population forecast given by President Lincoln in his 1862 Annual Message to Congress, which was based on simple geometric extrapolation, fits into this category.
Recently stochastic time series modelling is becoming more popular particularly within the framework of Box and Jenkins (1970), and Sabioa (1974) demonstrates this method applied to the population of Sweden.
Time series models, Land (1986) suggests, should only be considered for the short to medium term. Judgment

techniques should be used for long range forecasts.

The application of these methods has not been particularly successful as Lee (1974) comments: "not only have official US predictions been misleading with respect to future levels, they have also failed to anticipate the likely range of error".

Ascher (1979) illustrates that newer methods of total population forecasting are less successful than the early curve fitters such as Pearl (1925) for ten year forecasts and for five year forecasts the early forecasters were every bit as accurate as forecasts made via the elaborate procedures. Hajnal (1955) argues that population projections in the future as in the past will often be fairly wide of the mark - as often simple guesses would be "and that simple unpretentious short-term projections should be used to meet most practical needs for population forecasts." Hajnal further points out that, due to the slow rate at which population change takes place, usually almost any method of extrapolation from the past will give results that do not look absurd for a few years and "much of the elaborate technique of forecasters is expended in vain".

Dorn (1950) found that forecasts made in the USA over the period 1910 to 1950 by the complex component method gave predictions inferior to those made by extrapolation from the logistic model.

Apart from disappointing results of more complex forecasting, there are other problems that include the need for large amounts of data which is not always available, problems of updating as new information becomes available and that demographic theory as to how the components rates will change is inadequate.

Generally, the poor performance of the forecasting methods is blamed on the failure of human history to repeat itself and, as Hajnal (1955) comments, "the factors whose effects on future growth we can calculate are likely to be frequently outweighed by the unpredictable."

Murphy (1980) writes that, since forecasts have often been far from the path followed and so are misleading, there are two ways to overcome this problem, seeking a deeper insight into the process by considering more detailed information or, alternatively, to simplify the forecasting method. The stance taken in this Chapter will be the same as that followed by Murphy and the latter alternative is followed. (Land (1986) pointed out that the conventional response of increasing the complexity of models might well not be fruitful as small changes in such non-linear systems can cause major alterations. Hence there is a limit on accuracy.)

Using the simple trend time series models gives the advantages of:

(i) mechanical ease of producing forecasts;

(ii) forecasts are obtainable even with a very sparse data base;

(iii) forecasts and models are easily updated when fresh evidence becomes available.

Meade (1984), in examining the use of growth curves for forecasting market development, presents a list of criteria which growth curves must satisfy in order to be acceptable as forecasting methods. The criteria are:

(a) model validity - are there theoretical grounds for the choice of model?

(b) statistical validity - how well does the curve (model) fit past data?

(c) forecasting ability.

(d) interpretation of saturation level.

(e) forecast validity - can one derive a measure of the uncertainty of the forecast?

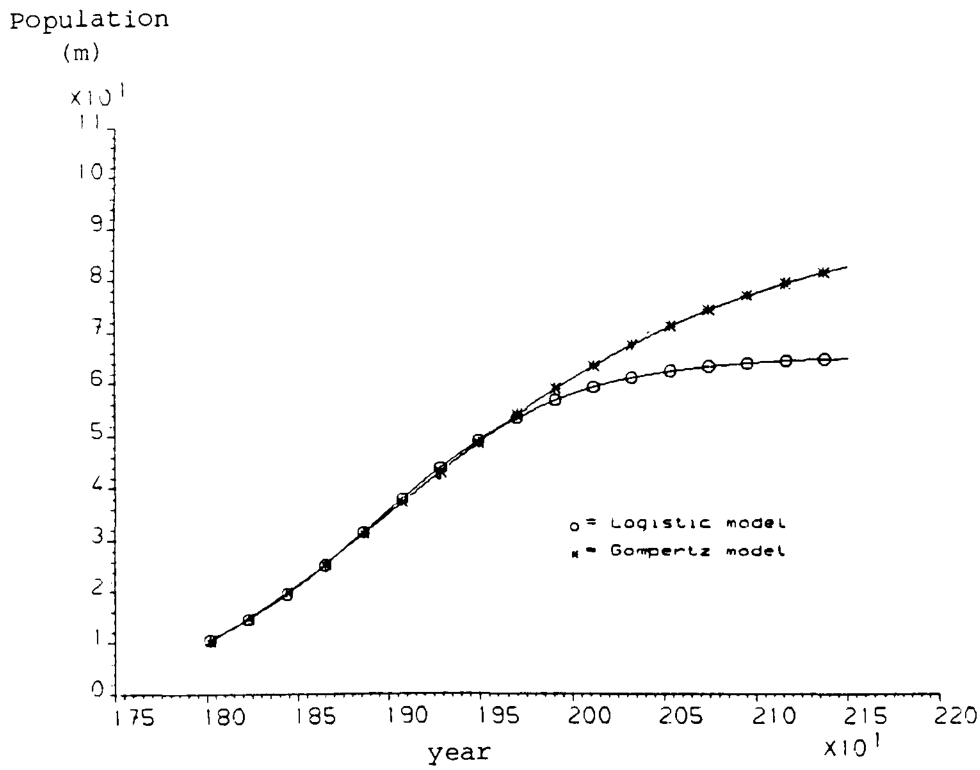
Statistical validity has been discussed in the previous Chapter and found to be satisfactory for sigmoid models. Model validity and the interpretation of saturation level are theoretical constructs and are discussed later in Chapter 5. Thus the subject of this Chapter will focus on forecasting ability and forecast validity.

After forming a basis on which forecast quality can be assessed in the next section, forecasts made by several trend models will be compared and their ability assessed in relation to future observations in section 3.3 and for different lead times in section 3.7. Section 3.4 examines the use of Box Jenkins ARIMA forecasts.

Section 3.5 considers reformulation of sigmoid models as a function of past observations rather than time. From this the state space approach is developed and the problem of forecast validity is addressed in section 3.6.

In the forthcoming section it is crucial to bear in mind that judgment plays a major role when forecasting using mathematical functions - not least in the choice of model. Unfortunately the nature of growth models gives little guidance as to the appropriate choice - many models fit past observations equally well, yet long-term extrapolation can be dramatically different. For example, Figure 3.1.1 displays the long-term forecasts of the 3-parameter Gompertz and logistic models that

Figure 3.1.1: 3-parameter logistic and Gompertz models fitted to the population of Great Britain 1801 to 1981 and extrapolated to the year 2151



fitted almost equally as well the population of Great Britain.

Therefore, population forecasting is not mere "mechanical mathematics" which can be labelled projection, but, like most forecasts, requires judgment and is as much an art as it is a science. These points are not pursued in this Chapter but are discussed again in Chapter 6 when the use of catastrophe theory (Zeeman 1976) is made to give a framework by which such unquantifiables can be examined.

In this Chapter total population only is considered and the methods developed are demonstrated with reference to the population of Great Britain, England and Wales, Scotland, Sweden and the USA.

3.2 Measure of Forecasting Ability

Various measures are available to assess the quality of a forecast when future observations become available. These measures include the mean absolute error (M.A.E.), root mean square error (R.M.S.E.), root mean square percentage error (R.M.S.P.E.) and the mean absolute percentage error (M.A.P.E.) - these are used and discussed by Ahlburg (1982) when considering forecasts of live births.

In this chapter M.A.P.E. is used which is given by:-

$$\frac{100}{n} \sum_{t=t_o+1}^{t_o+n} \left| \frac{P_t - \hat{P}_{t|t_o}}{P_t} \right|$$

This gives a measure of error independent of the magnitude of P_t (the population at time t) and will allow comparisons to be made across different populations.

3.3 Extrapolations from Simple Time Series Trend Models

In this section the simple trend models obtained in the previous Chapter are extrapolated to give the expected future value ($\hat{P}_{t|t_0}$) of total population. These forecasts are made from a base $t_L - i\Delta t$ where t is the census year, t_L being the latest census, Δt the interval between censuses and i = n, n-1, ..2,1. Forecasts are made up to the year t_L and compared to observed populations.

In section 2.7.1 of the previous Chapter predictions made by the simple models from the base year of 1930 were compared. From these Tables and a summary of them in Table 3.3.1 one can observe that generally sigmoid type models, although giving reasonable predictions (i.e. percentage error is less than 10% for periods of up to twenty years and in some cases up to fifty years) as in the case of the 4-parameter logistic extrapolated for the populations of Great Britain and Sweden, they are not consistently better than extrapolations of more simple models. The cubic, for example, was found in several cases to be superior (e.g. Sweden, see Table 3.3.2) and the linear model gave by far the closest predictions of the population of Great Britain over the period 1941 to 1981.

This seems to confirm Hajnal's (1955) view that simple models can give good forecasts. However, the problem faced by the forecaster is which simple model gives the most likely future path of population.

Table 3.3.1: M.A.P.E. of 50-year Forecasts

Population	3-Parameter Logistic	4-Parameter Logistic	3-Parameter Gompertz	4-Parameter Gompertz	Geometric	Modified Exponential	Linear	Quadratic	Cubic
Great Britain	7.62 Hi	5.09 Hi	11.93 Hi	8.80 Hi	40.83 Hi	19.98 Hi	1.77 Lo	17.17 Hi	9.00 Hi
England and Wales	7.19 Hi	4.04 Hi	10.17 Hi	8.04 Hi	40.91 Hi	20.28 Hi	3.02 Lo	16.76 Hi	8.68 Hi
Scotland	10.54 Hi	9.50 Hi	14.72 Hi	11.79 Hi	43.19 Hi	26.31 Hi	15.64 Hi	19.21 Hi	6.61 Hi
Sweden	5.20 Lo	6.10 Lo	3.47 Lo	5.71 Lo	11.51 Hi	3.05 Hi	4.90 Lo	2.06 Lo	6.28 Lo
U.S.A.	11.68 Lo	12.76 Lo	11.00 Hi	4.06 Hi	95.55 Hi	52.10 Hi	49.48 Lo	70.77 Lo	15.66
All 5 Populations	8.44	7.50	10.26	7.68	46.40	24.32	14.96	25.19	9.25

Hi = forecasts erring predominantly on the high side
Lo = forecasts erring predominantly on the low side

Table 3.3.2: Predictions of the Swedish Population M.A.P.E.

Model	Period of Prediction					
	1935-1981	1945-1981	1961-1981	1971-1981	1976-1981	1976-1981+
3-parameter logistic	5.20	7.53	5.42	3.35	2.51	3.51
4-parameter logistic	6.09	8.83	3.37*	1.54	0.68	1.28
3-parameter Gompertz	6.09	5.44	4.15	3.29	1.30	Fitting failed
Quadratic	2.06*	2.08*	4.16	2.45	3.79	0.45
Cubic	6.27	9.49	3.60	1.07*	0.55*	1.05*

Predictions were made from the models fitted from 1810 up to the base year of the prediction which, in this case, is five years less than the start of the prediction period, except for + which is predictions from models fitted over the period 1945 to 1975 to yearly estimates.

* = lowest M.A.P.E.

The cubic and other polynomial models are of particular interest to forecasters due in part to their ease of parameter estimation and construction of prediction intervals for forecasts. However, they have the disadvantages of tending to positive or negative infinity in the long-term, their parameters have little demographic interpretation and they tend to be strongly influenced by outliers and so fail on two of Meade's (1984) requirements, model validity and statistical validity. (This also applies to geometric models).

Table 3.3.3 restricts attention to the forecasts from logistic and Gompertz growth models over different lead times. The forecasts presented in Table 3.3.3 were derived from models fitted using proportionate weighting to give equal weighting to all data points - as described in Chapter 2, section 2.4. It could be argued that, in a population growth situation, models from unweighted fits would allow more recent points to carry more influence and, as a consequence, better forecasts would be produced. However, the results were inferior to those in Table 3.3.3 (the Scottish population is an exception to this - but not significantly so) so this was abandoned. Forecasts were improved by the addition of a percentage correction, so that the last fitted point corresponds to the last observation. This adjusted M.A.P.E is quoted in the adjusted column of Table 3.3.3.

From this Table one can draw the following general conclusions:

(i) all logistic models fail to give M.A.P.E.s of less than 10% for the USA (except the 3 parameter logistic fitted after escalation;

(ii) for the other populations it appears that models

Table 3.3.3: Comparison of M.A.P.E. for Logistic and Gompertz based forecasts
Population:- M.A.P.E.

Model	No.of yrs. forecasted	Sweden		Great Britain		England and Wales		Scotland		USA	
		Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
3PL	50	5.26 L	5.47 L	7.63 H	4.54 H	7.33 H	4.24 H	10.67 H	5.44 H	16.80 L	8.33 L
4PL		6.10 L	6.22 L	5.08 H	2.66 H	3.97 H	1.94 H	9.45 H	4.78 H	17.84 L	8.92 L
3PG		3.46 L	3.90 L	11.93 H	8.51 H	10.16 H	8.01 H	14.73 H	8.46 H	3.73 L	10.51 L
4PG		5.72 L	5.83 L	8.80 H	5.69 H	8.04 H	5.21 H	11.78 H	6.50 H	7.57 L	4.26 L
3PL	40	7.19 L	8.15 L	3.26 H	1.25 H	4.06 H	1.12 H	5.60 H	2.16 L	20.13 L	13.56 L
4PL		8.62 L	9.09 L	1.02 -	3.33 L	1.10 H	1.58 L	2.69 H	2.42 L	21.22 L	14.43 L
3PG		8.40 L	5.60 L	2.90 H	2.77 H	10.87 H	4.86 H	9.21 H	2.79 H	3.95 L	1.76 L
4PG		8.04 L	8.58 L	2.97 H	1.29 H	3.93 H	1.23 H	4.96 H	2.21 H	12.72 L	5.77 L
3PL	30	6.78 L	3.80 L	1.46 H	1.07 H	1.83 H	1.08 L	5.60 H	2.75 H	21.15 L	13.99 L
4PL		6.24 L	3.48 H	1.77 L	2.29 H	1.60 L	2.55 L	2.69 H	1.43 H	22.33 L	14.29 L
3PG		5.05 L	2.71 L	6.57 L	2.38 L	7.29 -	2.40 -	9.21 H	4.41 H	9.28 L	2.25 L
4PG		6.24 L	3.40 H	1.32 L	1.11 L	1.55 L	1.14 L	4.96 H	2.52 H	15.50 L	7.65 L
3PL	20	5.53 L	2.52 L	1.32 -	1.30 -	1.30 -	1.29 -	4.92 H	2.79 H	20.46 L	9.14 L
4PL		3.84 L	1.59 L	1.75 L	1.27 L	1.80 L	1.57 -	2.52 H	1.88 H	19.74 L	8.62 L
3PG		4.24 L	1.79 L	5.40 H	2.10 L	5.77 -	2.15 -	7.55 H	3.87 H	8.87 L	0.76 L
4PG		3.63 L	1.48 L	1.38 L	1.37 L	1.37 -	1.36 -	4.17 H	2.58 H	14.96 L	3.79 L
3PL	10	3.03 -	0.88 -	2.30 H	2.67 H	2.27 H	2.65 H	5.82 H	3.44 H	18.41 L	6.29 L
4PL		2.54 -	1.00 -	0.45 H	2.03 H	0.34 H	1.95 L	3.66 H	2.94 H	16.95 L	5.33 -
3PG		2.04 -	1.30 -	6.49 H	4.01 H	6.74 H	4.08 H	8.32 H	4.06 H	8.77 L	0.41 -
4PG		1.43 -	1.55 -	2.38 L	2.82 L	2.22 L	2.82 L	5.09 H	3.34 H	13.44 L	2.23 -
3PL*	10	-	-	-	-	-	-	2.68 H	0.71 H	1.78 L	1.60 L

PL = Parameter logistic model
PG = Parameter Gompertz model

L = Predominantly erring on low side
H = Predominantly erring on high side

* = Fit after escalation

which closely fit the population give good predictions with errors of less than 5% for a lead time of twenty years and often less than 10% for a lead time of fifty years;

(iii) the percentage correction only has significant beneficial effect for long lead times, being detrimental in the short term;

(iv) forecast inaccuracy generally increases as lead times lengthen, but not exponentially as Keyfitz (1979) suggested but much more gently.

(v) as a whole, the logistic models appear superior to the Gompertz models, the four-parameter models usually perform better.

This provides sufficient reasons for further investigation of sigmoids for forecasting as it would seem that, provided there are no phase changes, as described in Chapters 1 and 2, as was the case for the USA, one would anticipate useful predictions for lead times up to fifty years (which in practice is the maximum period that is required).

Before further examining time series forecasts made by the sigmoids, stochastic time series forecasting is considered with a view to arriving at an integrated forecasting methodology using both deterministic and stochastic modelling.

3.4 Stochastic Modelling of Population

Autoregressive Integrated Moving Average modelling is appealing due to the possibility of achieving great accuracy, can be applied to many time series requiring little previous assumptions to be made about the data and, more importantly for this investigation, confidence intervals can be generated automatically.

The procedure developed by Box and Jenkins (1970) is described by Montgomery (1976) and involves preliminary examination of a time series, model identification, parameter estimation and checking of the model form using diagnostic procedures before forecasts are made. Crucial to this is obtaining a stationary time series (constant variance throughout the series) which could be achieved if need be by differencing or by transformation of the series.

Saboia (1974) demonstrates the application of ARIMA modelling to the Swedish population. Saboia models the population of Sweden between 1780 to 1960 and generates forecasts of the population in 1965 and 1970 which he shows to be superior to those generated by the 3-parameter logistic and shows that the later observed populations lie within the 95% confidence limit.

The analysis carried out by Saboia is repeated to see if this result holds true for the 1975 and 1980 census count of Sweden.

ARIMA models were fitted to the population of Sweden between the years 1780-1960, 1780-1965, 1780-1970 and 1780-1980, and forecasts compared to observations and to those generated by the logistic models. The ARIMA modelling was carried out using the time series analysis routines in the NAG 10 Library incorporated into the programme PTRENDS (see Raeside, 1984). (The Time Series directives of the MINITAB statistical package were also found to be useful).

The choice of the ARIMA model was made by the following criteria:

(i) lowest sum of squares of residuals;

(ii) the autocorrelation of the residuals is not to be significant and, when summed to lag k, are less than $\chi^2_{k-p-q,\alpha}$ (i.e. the Box-Pierce Test, 1970), (p and q being the number of autoregressive and moving average terms respectively);

(iii) parameters of the model are significantly different from zero when compared to their standard errors;

(iv) the principle of parsimony - which is the ability to express the series by as few parameters as possible.

(v) how well the forecasts compare to observed values.

The models arrived at and their forecasts are presented in Table 3.4.1 along with associated 95% confidence intervals. Table 3.4.2 shows the results of ARIMA modelling applied to yearly estimates of the Swedish population between 1945 and 1975.

By reviewing this Table one can see that the stochastic ARIMA modelling does give good forecasts which have lower M.A.P.E.s than forecasts made by the deterministic logistic model. Also all observations lie within 95% confidence limits.

All the ARIMA models had to be differenced to achieve stationarity, an alternative is to remove the trend in the series. Thus, if a logistic model is used to remove this trend and ARMA models used to model the residuals, and forecasts from the ARMA part could be given confidence limits, a useful integrated

Table 3.4.1: ARIMA models of the population of Sweden and forecasts compared to logistic forecasts

Period 1780 - 1960

Models:

$$\text{ARIMA }(1,1,0)\quad P_t = 1.924P_{t-1} - \underset{(0.079)}{0.924}P_{t-2} + a_t$$

SSR = 0.210

$$\text{ARIMA }(0,2,1)\quad P_t = 2P_{t-1} - P_{t-2} + a_t - \underset{(0.131)}{0.680}a_{t-1}$$

SSR = 0.169

Forecasts:

Year	Observed	ARIMA (1,1,0)			ARIMA (0,2,1)			3-para logis.	4-para logis.
		Expected	Low	High	Expected	Low	High		
1965	7.767	7.735	7.584	7.890	7.733	7.595	7.871	7.632	7.610
1970	8.077	7.757	7.628	8.286	7.971	7.742	8.200	7.796	7.769
1975	8.208	8.163	7.591	8.697	8.209	7.887	8.532	7.957	7.924
1980	8.318	8.352	7.526	9.113	8.448	8.027	8.868	8.115	8.075
	M.A.P.E.	1.29			0.85			2.68	3.05

Period 1780 - 1965

Models:

$$\text{ARIMA }(1,1,0)\quad P_t = \underset{(0.0753)}{1.932}P_{t-1} - \underset{(0.0753)}{0.932}P_{t-2} + a_t$$

SSR = 0.211

$$\text{ARIMA }(0,2,1)\quad P_t = 2P_{t-1} - P_{t-2} + a_t - \underset{(0.131)}{0.661}a_{t-1}$$

SSR = 0.170

Forecasts:

Year	Observed	ARIMA (1,1,0)			ARIMA (0,2,1)			3-para logis.	4-para logis.
		Expected	Low	High	Expected	Low	High		
1970	8.077	8.021	7.871	8.171	8.018	7.882	8.155	7.930	7.926
1975	8.208	8.257	7.931	8.583	8.269	8.041	8.498	8.102	8.108
1980	8.318	8.477	7.945	9.009	8.520	8.196	8.844	8.271	8.288
	M.A.P.E.	1.09			1.32			1.22	1.15

Period 1780 - 1970

Models:

$$\text{ARIMA }(1,1,0)\quad P_t = \underset{(0.072)}{1.946}P_{t-1} - \underset{(0.072)}{0.946}P_{t-2} + a_t$$

SSR = 0.214

$$\text{ARIMA }(0,2,1)\quad P_t = 2P_{t-1} - P_{t-2} + a_t - \underset{(0.134)}{0.632}a_{t-1}$$

SSR = 0.173

Forecasts:

Year	Observed	ARIMA (1,1,0)			ARIMA (0,2,1)			3-para logis.	4-para logis.
		Expected	Low	High	Expected	Low	High		
1975	8.208	8.370	8.221	8.519	8.351	8.215	8.487	7.930	7.926
1980	8.318	8.648	8.322	8.974	8.626	8.395	8.856	8.102	8.108
	M.A.P.E.	1.01			2.72				

Period 1780 - 1975

Models:

$$\text{ARIMA }(1,1,0)\quad P_t = \underset{(0.0721)}{1.904}P_{t-1} - \underset{(0.0721)}{0.904}P_{t-2} + a_t$$

SSR = 0.238

$$\text{ARIMA }(0,2,1)\quad P_t = 2P_{t-1} - P_{t-2} + a_t - \underset{(0.115)}{0.714}a_{t-1}$$

SSR = 0.192

Forecasts:

Year	Observed	ARIMA (1,1,0)			ARIMA (0,2,1)			3-para logis.	4-para logis.
		Expected	Low	High	Expected	Low	High		
1980	8.318	8.264	8.171	8.482	8.434	8.292	8.575	8.172	8.247
	M.A.P.E.	0.65			1.40			1.76	0.85

The figures in brackets are the standard error of the parameters.

Table 3.4.2: Models fitted to yearly counts of the Swedish population

Year Fore-casted	Observed	ARIMA (1,1,0)*			ARIMA (0,1,2)**			ARIMA (1,1,1)***			3-parameter logistic	4-parameter logistic
		Fore-cast	Low	High	Fore-cast	Low	High	Fore-cast	Low	High		
1976	8.244	8.244	8.234	8.253	8.247	8.238	8.256	8.241	8.232	8.250	8.284	8.286
1977	8.267	8.284	8.264	8.303	8.292	8.272	8.312	8.282	8.262	8.303	8.327	8.330
1978	8.284	8.326	8.297	8.356	8.344	8.313	8.375	8.328	8.296	8.360	8.369	8.373
1979	8.303	8.371	8.331	8.411	8.396	8.357	8.435	8.377	8.335	8.419	8.414	8.415
1980	8.318	8.418	8.368	8.468	8.448	8.402	8.494	8.427	8.375	8.478	8.452	8.458
1981	8.323	8.466	8.406	8.526	8.500	8.449	8.552	8.478	8.418	8.597	8.492	8.499
	M.A.P.E.	0.74			0.98			0.80			1.20	1.25

* ARIMA (1,1,0) → $z_t = z_{t-1} + \underset{(0.150)}{0.772}z_{t-1} + a_{t-1}$

sum of squares = 0.0055

** ARIMA (0,1,2) → $z_t = z_{t-1} - \underset{(0.1577)}{1.1203}a_{t-1} - \underset{(0.1641)}{0.5986}a_{t-2}$

sum of squares = 0.0045

*** ARIMA (1,1,1) → $z_t = z_{t-1} + \underset{(0.220)}{0.588}z_{t-1} + a_t - \underset{(0.220)}{0.483}a_{t-1}$

sum of squares = 0.0048

The figures in brackets are the standard errors of the parameters.

3-parameter logistic

$$N_t = \frac{10.822}{1 + \left[\frac{10.822}{0.236} - 1\right]e^{-0.0221t}}$$

sum of squares = 0.0006

4-parameter logistic

$$N_t = \frac{146.9758}{1 + \left[\frac{146.9758}{114.3296} - 1\right]e^{-0.0095t}} + (-133.9449)$$

sum of squares = 0.0005

forecasting method would result.

Thus the model would be:

$$P_t = \frac{N_\infty}{1 + \left[\frac{N_\infty}{N_o} - 1\right] e^{-ct}} + d \qquad \text{(deterministic part)}$$

$$+ \phi_1 e_{t-1} + \phi_2 e_{t-2} \cdots \phi_p e_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta_q a_{r-q}$$

(stochastic part)

This is appealing as it would allow the cyclical nature of the residuals that was noted when the models were fitted in Chapter 2 (section 2.4) to be "carried" forward in the forecasts.

This "two stage" modelling was applied to the Swedish population between 1780 and 1960 and forecasts compared to observed population over the period 1965 to 1980. The best model that was derived was a three-parameter logistic with an ARIMA (2,0,0) model of the errors. This is parameterised as:

$$P_t = \frac{18.924}{1 + \left[\frac{18.924}{1.513} - 1\right] e^{-0.0096t*}} + \underset{(0.163)}{1.130} e_{t-1} - \underset{(0.165)}{0.368} e_{t-2} + a_t$$

(standard errors in brackets)

$t* = \text{year} - 1750$

This performs very well when compared to future observations as can be observed from Figure 3.4.1 having a M.A.P.E. of 1.46 which is much lower than the 4-parameter logistic model on its own. However, it is the generation of confidence limits that gives this approach the advantage.

Applying this to the population of Great Britain and component populations gives the models displayed in Table 3.4.3. As can be observed from examining Table 3.4.4, forecasts from these compare closely, except for Scotland, with O.P.C.S. (1984) projections made by the more complex component method.

Figure 3.4.1: 1960 based forecasts of the Swedish population

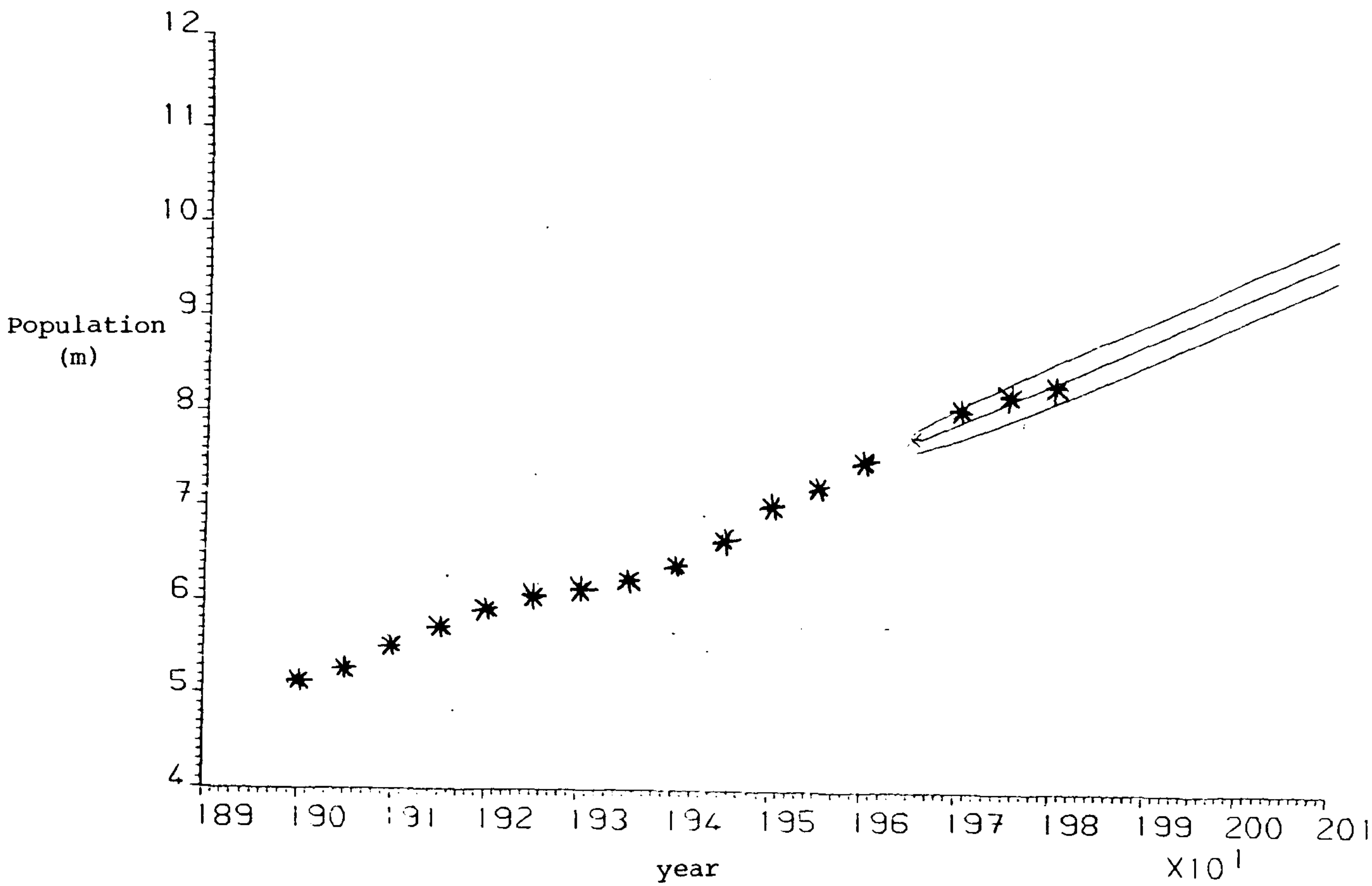


Table 3.4.3: Models for forecasting the populations of Great Britain, England and Wales and Scotland

Great Britain

$$P_t = \frac{57.376}{1 + \left[\frac{57.376}{7.241} - 1\right]e^{-0.022t^*}} + 3.321 + \text{ARIMA}\ (3,0,0)$$

$$\text{where ARIMA}\ (3,0,0) = \underset{(0.247)}{0.264}e_{t-1} - \underset{(0.115)}{0.316}e_{t-2} + \underset{(0.277)}{0.778}e_{t-3} + a_t$$

England and Wales

$$P_t = \frac{52.584}{1 + \left[\frac{52.584}{6.152} - 1\right]e^{-0.022t^*}} + 2.787 + \text{ARIMA}\ (2,0,1)$$

$$\text{where ARIMA}\ (2,0,1) = \underset{(0.200)}{-1.056}e_{t-1} + \underset{(0.198)}{0.621}e_{t-2} + a_t - \underset{(0.359)}{0.902}a_{t-1}$$

Scotland$^+$

$$P_t = \frac{5.28}{1 + \left[\frac{5.28}{0.65} - 1\right]e^{-0.035t^*}} + \text{ARIMA}\ (2,0,0)$$

$$\text{where ARIMA}\ (2,0,0) = \underset{(0.250)}{-1.046}e_{t-1} + \underset{(0.251)}{0.542}e_{t-2} + a_t$$

t^* = year - 1801

$^+$ = fitted to yearly estimates

Standard errors in brackets.

Table 3.4.4: Forecasts compared

Year	Great Britain		England & Wales		Scotland	
	OPCS	Logistic Range	OPCS	Logistic Range	OPCS	Logistic Range
1982	54.729	54.339-54.511	49.327	49.211-49.348	5.136	5.132-5.159
1983	54.726	54.390-54.733	49.604	49.264-49.538	5.122	5.103-5.184
1984	54.745	54.439-54.954	49.635	49.316-49.727	5.110	5.095-5.200
1985	54.783	54.486-55.172	49.684	49.366-49.913	5.099	5.090-5.217
1986	54.851	54.530-55.388	49.759	49.413-50.097	5.092	5.089-5.234
1991	55.386	54-722-56.439	50.314	49.622-50.990	5.072	5.120-5.300
1996	56.002	54.788-56.542	50.944	50.114-51.486	5.058	5.136-5.335
2001	56.624	54.808-56.599	51.416	50.563-51.939	5.108	5.150-5.349
2011	56.886	56.196-58.017	52.011	51.258-52.809	4.875	5.157-5.356
2021	57.796	56.561-58.722	53.042	51.786-53.570	4.754	5.161-5.361

The OPCS projections are 1981 based.

Although the main concern of this Chapter is forecasting, the models can also be used to generate backcasts. This has been done for Sweden and for England.

For Sweden the model:

$$P_t = \frac{13.212}{1 + \left[\frac{13.212}{1.305} - 1\right] e^{-0.0118t*}} + \text{ARIMA } (2,0,0)$$

where t* = year - 1750

and ARIMA (2,0,0) = $\underset{(0.166)}{-1.160}e_{t-1} + \underset{(0.174)}{0.417}e_{t-2} + a_t$

was used to give backcasts prior to 1810, which are presented in Table 3.4.5.

Table 3.4.5: Backcasts of Sweden compared to observations

Year	Observed Value	Forecast Value	Percentage Error	50% Confidence Limits Lower - Upper bound
1805	2.427	2.358	2.8	1.986 - 2.644
1800	2.347	2.254	4.0	1.857 - 2.605
1795	2.281	2.154	5.6	1.763 - 2.534
1790	2.188	2.058	5.9	1.679 - 2.453
1785	2.150	1.966	8.6	1.596 - 2.371
1780	2.118	1.877	11.4	1.508 - 2.292
1775	2.021	1.792	11.3	1.417 - 2.217
1770	2.043	1.710	16.3	1.324 - 2.144
1760	1.924	1.556	19.0	1.147 - 1.999
1750	1.781	1.415	20.6	0.988 - 1.855

From the figures presented in the above Table one can see that, although the forecasts are consistently lower than the observed values, they do not deviate by more than 10% until 1780 which is a run of 25 years. Also all observations are contained within 50% confidence limits.

For England backcasts were made by the model:

$$P_t = \frac{50.247}{1 + \left[\frac{50.247}{5.998} - 1\right] e^{-0.022t^*}} + 2.326 + \text{ARIMA}\ (2,0,1)$$

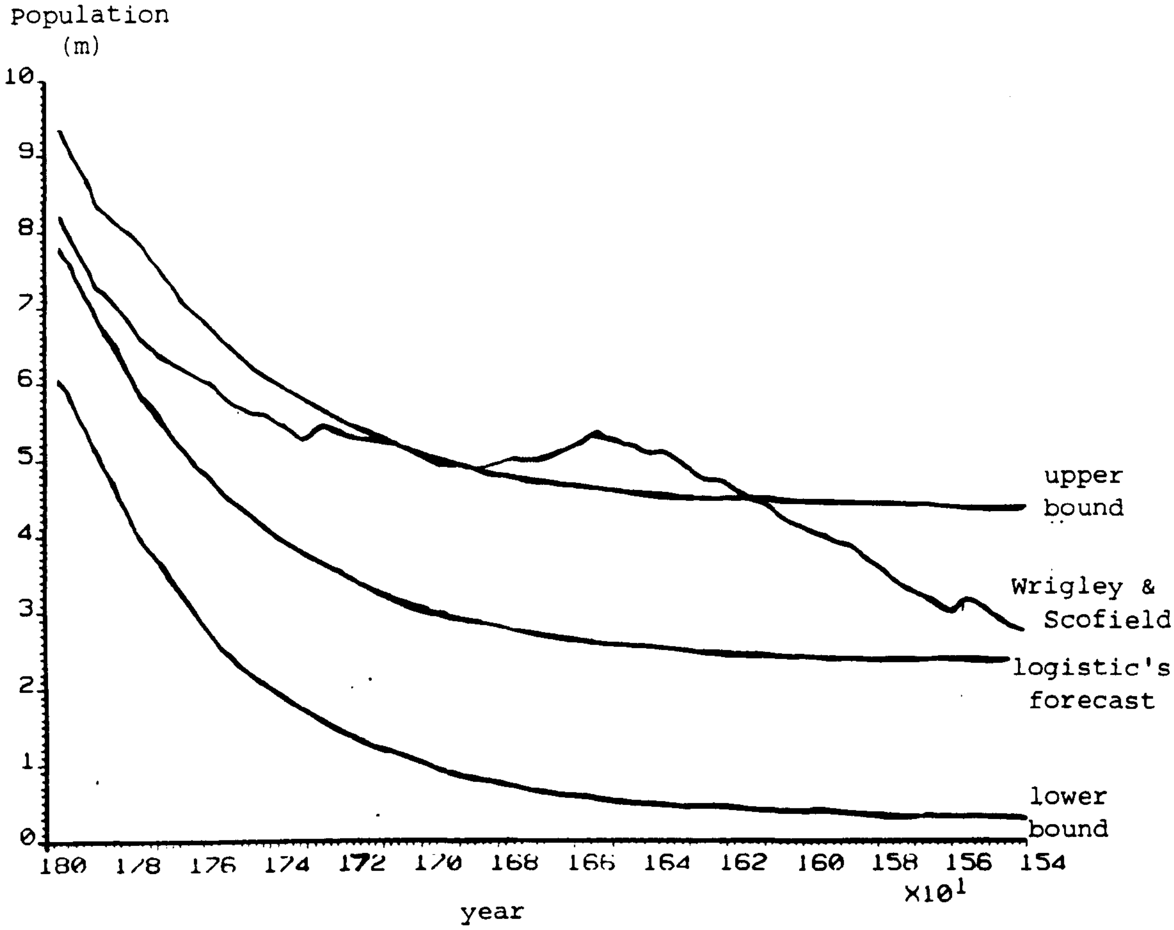
where t^* = year - 1801 and

$$\text{ARIMA}\ (2,0,1) = \underset{(0.193)}{-1.085e_{t-1}} + \underset{(0.188)}{0.723e_{t-2}} + a_t + \underset{(0.283)}{0.990a_{t-1}}$$

These are compared to the reconstruction of the English population given by Wrigley and Scofield (1981) in Figure 3.4.2. This Figure illustrates that the backcasting performance of the above model is not particularly good. This may be due to a change in the phase of growth as the English population benefited from agricultural improvements and industrialisation and urbanisation began. (This is discussed further in Chapter 5.)

However, there are major problems with this approach to total population modelling and forecasting which makes this an undesirable method. The most notable problems are:

Figure 3.4.2: Back projections for the English population



(a) A great deal of data is required. Box and Jenkins recommend at least 50 observations are required. Saboia (1977) does not see lack of data as such a problem, arguing that patterns of growth fifty years ago have little to do with current patterns. One might think that, if the ARIMA models are kept parsimonious, problems of model specification might be lessened. But for many populations fewer than twenty observations are available, and so it seems inevitable that parameter estimation for the ARIMA models is uncertain and forecasts and confidence limits must be of doubtful validity.

(b) The method is a slow means of model building (even with computer packages) and, since there are many choices of model, one may reject or not even consider a good model accepting instead a poor model which happens to pass a statistical test, or to agree with preconceptions. Biases and other prejudices, especially in the identification phase, can easily arise as has been highlighted by Pittenger (1978).

(c) Eventual forecasts of ARIMA models tend to either constant levels or to converge to constant oscillations, and so it is only really a short-term forecasting procedure.

(d) Updating, when new information becomes available, requires remodelling the whole series, unless recursive methods are employed.

(e) Confidence intervals are wide (for example, Land (1986) comments that the 95% confidence interval after five years could be as wide as ±30% of the forecast magnitude and within fifteen years this could be as large as ±50%).

Thus, for total populations ARIMA modelling is considered to be inadvisable, mainly due to shortages of data, and fails Meade's (1984) criteria on grounds of statistical validity.

3.5 "Recursive" Forecasting

Extrapolation of the models, as in section 3.3, is problematic if the extrapolations are to be used as forecasts. Problems which can be identified are:

(i) assuming constant model parameters;

(ii) difficulty in obtaining statistically valid confidence intervals (this will be addressed in the next section);

(iii) problem of identifying phase changes and altering forecasts to accommodate such a change.

A recursive estimation procedure, in which the parameters that are used to give estimates of the next period in the future are estimated from past knowledge and updated by the most recent observation and, if the variance of the error at each fit can be extrapolated, might offer a means by which these problems can be tackled.

In this procedure the population model is no longer specified as a function of time, but as a function of past observations, i.e:

$$P_t = g(\beta_0, \beta_1, \ldots\ldots, P_{t-1}, P_{t-2}, \ldots\ldots) + \text{error}$$

Various methods are available by which this can be carried out such as using Discounted Least Squares (D.L.S.) (see Montgomery (1976) for details, and Young and Ord (1984) for application to sigmoids). However, in D.L.S. the use of a smoothing parameter is critical, for which little guidance can be given and is often chosen arbitrarily.

Murphy (1980) illustrates the use of simple recurrence re-expression of the trend models and shows the generation of forecasts by the appropriate recurrence relationship. For example, Murphy writes the 3-parameter logistic model as:

$$P_t = \frac{1/P_{t-2} - 1/P_{t-3}}{1/P^2_{t-1} + 1/P^2_{t-2} - 1/P_{t-1}P_{t-2} - 1/P_{t-1}P_{t-3}}\ .$$

Murphy examines the forecasting performance of geometric, modified exponential, linear and logistic models. (Constant and cyclical models are also examined; these may have application in cases where the population has entered a stationary state.) Murphy applies the models to the populations of Great Britain, France, USA and Sweden. The logistic model shown above was found to give unsatisfactory forecasts (i.e. infinite or negative values). (The cyclical models were found unsatisfactory.)

Murphy introduces an amended logistic to produce bounded forecasts and which is capable of handling declining or fluctuating populations by defining the upper asymptote in terms of recent observations. This is defined as $P_2(P_2/P_1)^6$, the value to which the population would grow in approximately two generations at the current rate of increase. Thus the 3-parameter logistic is reformulated as:

$$P_t = \frac{a}{1 + \left[\frac{a}{P_{t-1}} - 1\right]^2 / \left[\frac{a}{P_{t-2}} - 1\right]}, \text{ where } a = P_2(P_2/P_1)^n \quad (n = 6)$$

Murphy found that this form gave satisfactory forecasts (although the simple geometric model also performed well), and betters the performance of official component based forecasts.

This amended form of the logistic, which Murphy refers to as an asymptotic model, has the interesting property of updating itself as new information becomes available and should "follow" escalations and de-escalations as forecasts are generated from only the two most recent values.

Table 3.5.1 displays the M.A.P.E.s of forecasts of the populations considered in this Chapter for different base years up to the last available censuses.

Table 3.5.1: M.A.P.E. of forecasts made by the "asymptotic" model

Base Year and (lead time)	M.A.P.E. Great Britain	England & Wales	Scotland	Base Year and (lead time)	M.A.P.E. Sweden	U.S.A.
1881(100)	4.42 Hi	3.63 Lo	11.36 Hi	1880(100)	11.41 Lo	6.49 Lo
1891 (90)	1.84 -	1.48 -	5.80 Hi	1890 (90)	17.05 Lo	5.77 Lo
1901 (80)	9.80 Hi	8.08 Hi	23.69 Hi	1900 (80)	6.62 Lo	5.41 -
1911 (70)	10.43 Hi	10.16 Hi	12.51 Hi	1910 (70)	5.89 Lo	5.49 -
1921 (60)	3.68 Lo	4.48 Lo	2.81 Hi	1920 (60)	7.01 Lo	7.98 Lo
1931 (50)	2.88 Lo	2.05 Lo	6.38 Hi	1930 (50)	11.69 Lo	4.83 -
1941 (40)	7.06 Lo	3.16 Lo	7.43 Hi	1940 (40)	11.11 Lo	19.01 -
1951 (30)	1.31 -	1.68 Lo	7.71 Hi	1950 (30)	2.49 Hi	6.05 -
1961 (20)	1.57 -	1.59 -	2.01 Hi	1960 (20)	0.93 Lo	4.59 Hi
1971 (10)	3.99 Hi	4.15 Hi	2.76 Hi	1970 (10)	2.65 Hi	0.50 Hi

Hi = forecasts are predominantly erring on high side
Lo = forecasts are predominantly erring on low side

As can be observed from this Table, forecasts made by this simple method are good, and forecasts over the last fifty years compare favourably with those made by extrapolation, the models fitted as functions of time (see Table 3.3.3). The low error for the USA and Scotland after 1961 indicates that this simple method has been able to adjust to the phase change which occurred in these nations.

Comparing the average M.A.P.E. for each population by forecasts made in the last fifty years gives Table 3.5.2.

Table 3.5.2: Average M.A.P.E. for forecasts over the last 50 years

Population	Method: M.A.P.E.				
	Asymptotic	Period Models		Kalman Filter	
		3-parameter logistic	4-parameter logistic	Logistic	Exponential
Gt.Britain	3.36	2.60	2.02	2.05	3.51
England & Wales	2.53	3.34	1.78	1.66	4.23
Scotland	5.26	6.36	4.14	2.56	3.51
Sweden	5.77	3.98	4.57	1.39	2.43
U.S.A.	7.00	15.01	15.98	8.12	3.21
All five popns.	4.78	6.26	5.70	3.16	3.38

By considering Table 3.5.2 there seems to be only a slight advantage in using the 4-parameter time based logistic to the simple asymptotic model, and this advantage is lost if there has been a major phase change as with the USA (the 3-parameter logistic model being inferior to the 4-parameter model.)

Thus it is worthwhile to investigate further the simple type of recursive forecasting. However, for this method it is not possible to obtain measures of uncertainty (confidence intervals) and so this type of forecasting will not be pursued but will be used as a benchmark for assessing the worth of more complex forecasting procedures.

Meade (1985) illustrates the merits of Kalman filtering (for details, see Harrison and Stevens, 1976) for generating forecasts of popular record sales.

This method is based on the recursive estimation of parameters using the relationship:

$$P_t = h_t(\beta_t) + e_t.$$

The 3-parameter logistic is reformulated from

$$P_t = \frac{N_\infty}{1 + \left(\frac{N_\infty}{N_o} - 1\right)e^{-ct}} \quad \text{to} \quad P_t = \frac{P_{t-1}}{\beta_1 + \beta_2 P_{t-1}}$$

where $\beta_1 = e^{-c}$ and $\beta_2 = \frac{1 - e^{-c}}{N_\infty}$

The extended Kalman filter is used to update the estimates of the coefficients β_1 and β_2 each time a new observation becomes available, i.e.

$$\hat{\underline{\beta}}_t = \hat{\underline{\beta}}_{t-1} + \underline{W}_t \text{ ; } \underline{W}_t \text{ is a "correction" term}$$

$$P_t = H_t'\hat{\underline{\beta}}_t + e_t \text{ where } H_t'$$

is the transpose of the matrix H_t of partial derivatives of P_t, i.e. $H_t = \left.\frac{\partial h_t(\beta)}{\partial \beta}\right| \beta = \hat{\beta}_{t|t+1}$ and e_t is a Gaussian error term.

In the procedure described by Meade (1985) the covariance matrix of the coefficient estimates is:

$$\Sigma_{t|t-1} = E\{[\beta_t - \hat{\beta}_{t|t-1}][\beta_t - \hat{\beta}_{t|t-1}]'\}$$

When a new observation becomes available the updating equations are:

$$L_t = \Sigma_{t|t-1}H_t(H_t'\,\Sigma_{t|t-1}H_t + R_t)^{-1}$$

$$\hat{\underline{\beta}}_{t+1|t} = \hat{\underline{\beta}}_{t|t-1} + L_t\left|P_t - H_t'\,\hat{\beta}_{t|t-1}\right|$$

$$\Sigma_{t|t} = \Sigma_{t|t-1} - L_tH_t'\Sigma_{t|t-1}$$

$$\Sigma_{t+1|t} = \Sigma_{t|t} + Q_t$$

where R_t is the variance of e_t and is estimated by:

$$S(\hat{\beta}) = \frac{1}{n-2}\sum_{t=1}^{n}[P_t - h_t(\hat{\beta})]^2$$

This procedure requires initialisation by providing initial estimates of $\underline{\beta}_o$, R_o and Σ_o. $\underline{\beta}_o$ and R_o are estimated from the application of the NAG Library non-linear regression routine EØ4FDF to fit the model $P_t = P_{t-1}/(\beta_1+\beta_2P_{t-1})$ to an initial portion of the P_t series.

Σ_o is estimated from $2S(\hat{\beta})G^{-1}$ where $G = J'J$, J being the Jacobian matrix:

$$J_{ij} = \frac{\partial ht}{\partial \beta_j}(\beta).$$

This method was implemented to give forecasts of the populations used in this Chapter. (The computer program that was used is described in Raeside, 1985).

M.A.P.E.s of forecasts made by this method compared to future observations, up to the most recent census count, are given in Table 3.5.3. The results are encouraging and display very satisfactory forecasting behaviour, even for very long lead times; forecasts for Sweden are particularly good. The USA is an exception to this although, as can be seen from Table 3.5.2, the Kalman filter is a big improvement on straightforward extrapolation from period based models, indicating updating is moving forecasts in the correct direction to accommodate the phase change, but still does not better the simple asymptotic model. Also the Kalman filter approach gives the lowest M.A.P.E. for all countries' forecasts combined over the last fifty years.

The case of Sweden is particularly encouraging as a forecast made in 1900 gives remarkably close predictions for a period of about one hundred years see Figure 3.5.1.

The poor quality of post-1950 forecasts for the USA indicates that the weight of the new update is insufficient to cause the filter to change rapidly enough to allow for the escalation of the USA population prior to 1950 (although for Scotland there are signs that the filter is adjusting satisfactorily).

Since escalation has been judged a radical departure from the expected path, then the update of the new parameter should not be based on all previous estimates but only the most recent observations.

To accomplish this the following steps were built into the

Table 3.5.3: Kalman filter logistic forecasts

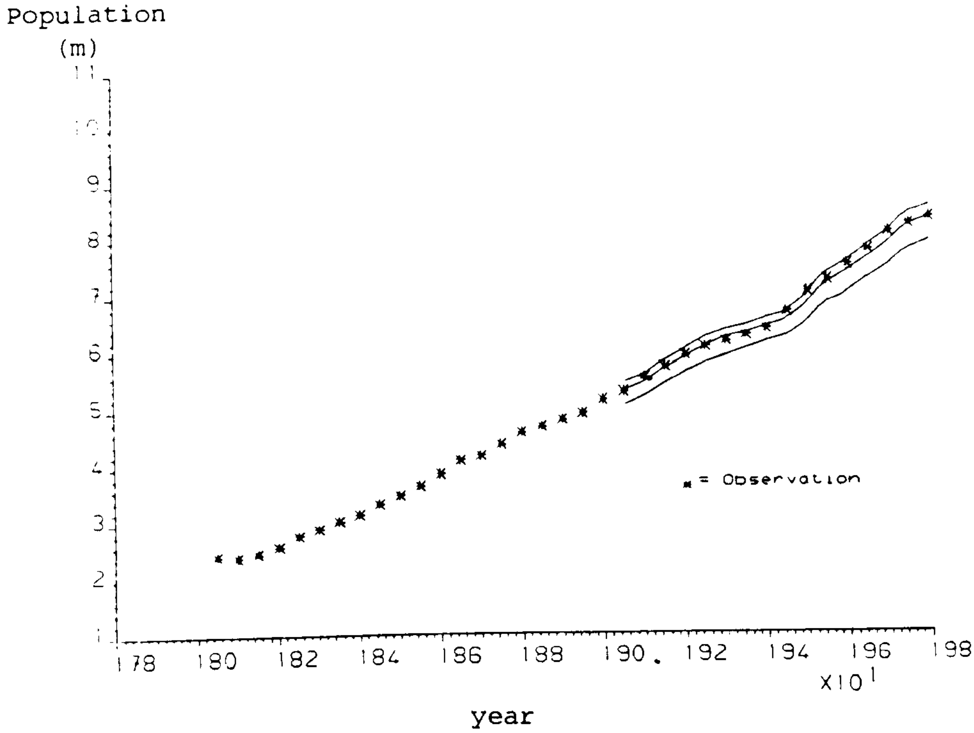
Base Year of forecast	Great Britain	England & Wales	Scotland	Base Year of forecast	Sweden	U.S.A.
1881	2.99 Hi	4.42	3.46 Hi	1880	1.55 -	9.06 Lo
1891	3.25 Hi	3.75	3.54 Hi	1890	1.40 -	5.01 Lo
1901	3.81 Hi	4.48	7.00 Hi	1900	1.27 -	7.04 Lo
1911	3.84 Hi	4.67	4.76 Hi	1910	1.02 -	5.13 Lo
1921	1.61 -	1.16	3.96 Hi	1920	1.16 -	7.43 Lo
1931	1.30 -	1.25	2.62 Hi	1930	1.44 -	6.89 Lo
1941	1.57 -	1.26	1.91 Hi	1940	1.98 Lo	10.11 Lo
1951	3.29 -	1.66	2.03 Hi	1950	0.85 Lo	7.01 Lo
1961	1.85 -	1.91	2.50 Hi	1960	1.07 -	8.39 Lo
1971	2.23 Hi	2.24	3.72 Hi	1970	1.63 -	8.22 Lo

Hi = forecast is predominantly erring on the high side
Lo = forecast is predominantly erring on the low side

The first eight points of each series is used for initialisation.

Figure 3.5.1: 1900 based forecasts of Swedish population

(the solid line represents the forecasts)



updating procedure:

If the ratio $R_t/R_{t-1} > 2$, then

$$\hat{\beta}_1 t/_{t+1} = \frac{1/P_t - 1/P_{t-1}}{1/P_{t-1} - 1/P_{t-2}}$$

and $$\hat{\beta}_2 t/_{t+1} = 1/P_t - \frac{\hat{\beta}_1 t/t+1}{P_{t-1}}$$

Meade (1985) suggests altering Q_t to a significant size in the updating procedure when a major change occurs, but, since this is an arbitrary procedure for which no guidance is given, the above procedure is preferred.

However, when this is applied to the filtering of the USA, a substantial improvement is not observed (nor was there an improvement when Q_t was set to a significant size).

The exponential model formulated as $P_t = \beta_1 + \beta_2 P_{t-2}$ and the Gompertz model formulated as $P_t = \beta_2/\beta_1 P_{t-1}$ were also examined in this way. It was only for the USA that the exponential performed better than the logistic, as can be seen from Table 3.5.2. The Gompertz model failed to be successfully updated as its parameters became non-positive. Therefore, the Gompertz and exponential models are given no further consideration, although it should be noted that the Kalman filter implementation of the exponential outperforms period logistic models. Thus, Kalman filtering is preferred to period based forecasting.

Apart from the quality of forecasts, an approach such as Kalman filtering is appealing in demographic forecasting as forecasts are made conditional principally on the most recent observations. Hence, the inherent serial correlation in the series can be used to advantage in making forecasts.

3.6 Measures of Uncertainty on Forecasts

It is important to get some notion of the degree of uncertainty that is associated with forecasts. In component-cohort based projections, high and low projectons are given which are often taken by users as the likely range in which the future will lie. However, these "bounds" have no statistical validity, being merely judgmental extrapolations.

Leach (1981) demonstrates the calculation of variant projections for the logistic, but again these are rather arbitrarily defined and his procedure is not easily generalised.

The ARIMA modelling gave statistical confidence limits but the method is not suitable for widespread use and is in parts dubious. Brand et al. (1973) derive statistical confidence limits for the logistic but the method is cumbersome, and Meade (1984) suggests a more generalised and more easily implemented approach by using Monte Carlo simulation methods. The coefficients of the period models could be sampled from a P variate normal distribution, using the parameter estimates and covariance matrix and an error term superimposed. But, for sigmoid models used in relation to population, this method may not be suitable since the calculation of the covariance matrix that adequately accounts for inherent correlation in the data series is not certain, and one must be cautious about relying on the properties of the asymptotic covariance matrix as the models are non-linear (see Section 2.6). This is discussed further in Chapter 2, Section 2.4.

A major advantage of using Kalman filtering is that estimates of the covariance matrix, and the error variance as well as the parameter estimates which are computed, produce sufficient information for a Monte Carlo simulation as described above. By generating sufficient simulations of the series (1000 in this implementation), then confidence intervals can be generated by the precentile method.

Meade (1985) describes the method in relation to Kalman filtering. To obtain future realisation of the series, values of $\hat{\beta}_{1t+1}$ and $\hat{\beta}_{2t+1}$ are sampled from the bivariate normal distribution:

$$\hat{\underline{\beta}}_{t+1} \sim N(\hat{\underline{\beta}}_{t+1/t}; \Sigma_{t+1/t}).$$

To carry this out the NAG Library routine GØ5EZF was used. P'_{t+i} is then computed from $h_{t+i}(\beta'_{t+1}) + e_{t+i}$, where e_{t+i} is sampled from a Gaussian distribution with zero mean and variance R_{t+1}, and are generated using the NAG Library routine GØ5DDF.

The precentile confidence limits are obtained by using the routine of Buckland (1985).

Forecasts with 95% confidence limits are illustrated for the populations for forecasts made in 1930 by Kalman filter of the logistic over the period up to 1980 in Figures 3.6.1 - 3.6.5. These appear satisfactory and one can conclude that, if the Kalman filter approach is used, then the logistic model satisfies the forecasting criteria of Meade (1984) given in the introduction.

The Kalman filter approach has a further merit in that it can be made to adjust rapidly for cases where there have been a change in the phase of growth. However, this adjustment can only take place after the event and was not found to be sufficient for the USA suggesting the rule that, if a change of phase is detected, then one should switch to a simple method such as Murphy's (1980) asymptotic model which only uses localised information.

The length of lead time of these forecasts is very much

Figure 3.6.1: 1931 based forecasts of Great British population

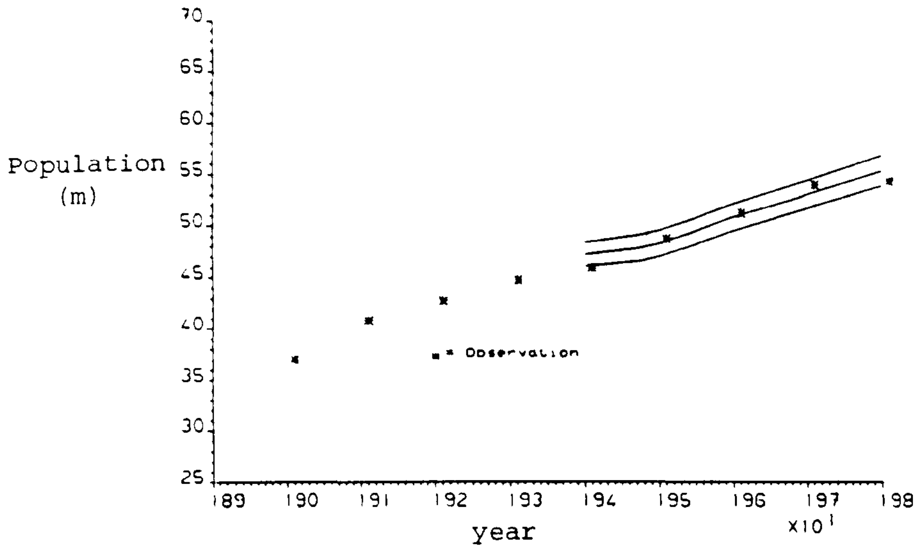


Figure 3.6.2: 1930 based forecasts of Swedish population

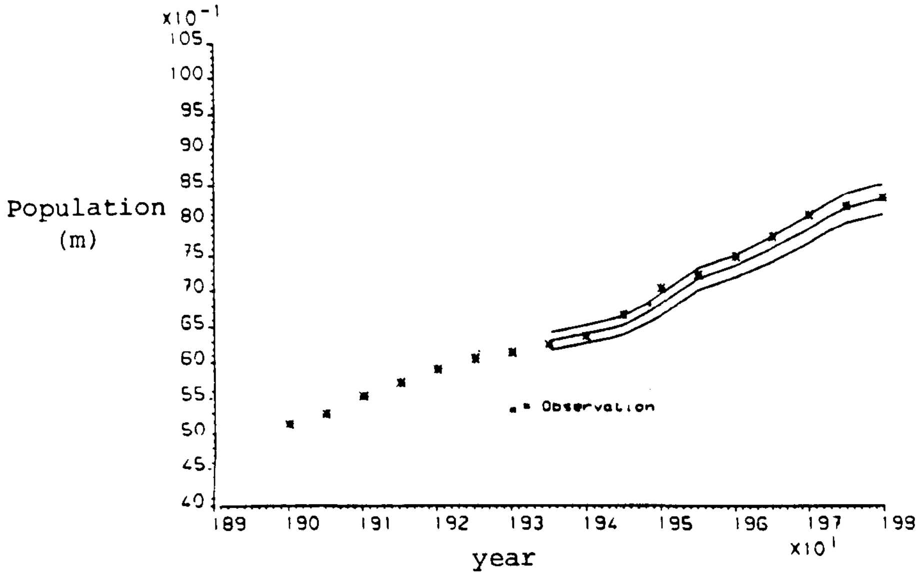
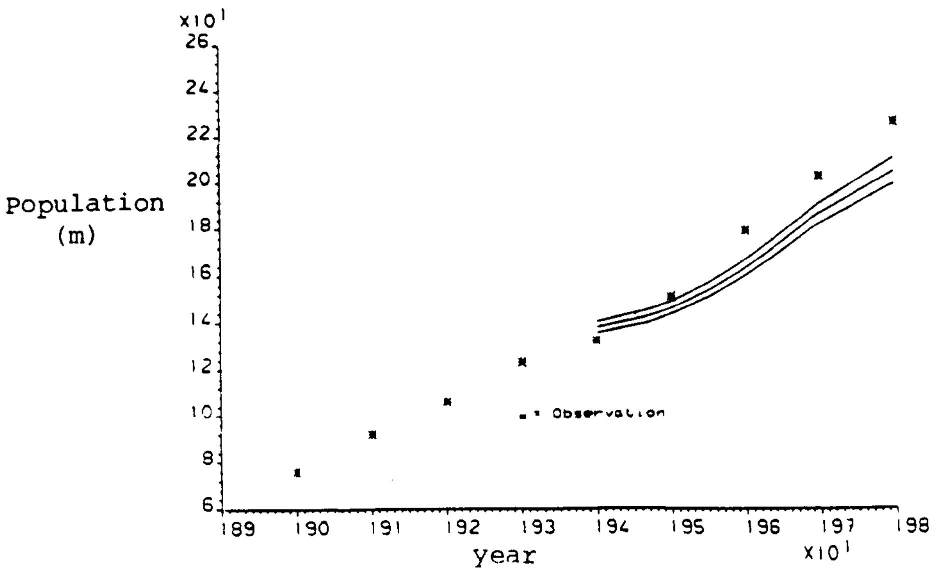


Figure 3.6.3: 1930 based forecasts of USA population



Forecasts and the 95% confidence intervals are the solid lines.

Figure 3.6.4: 1931 based forecasts of the English and Welsh population

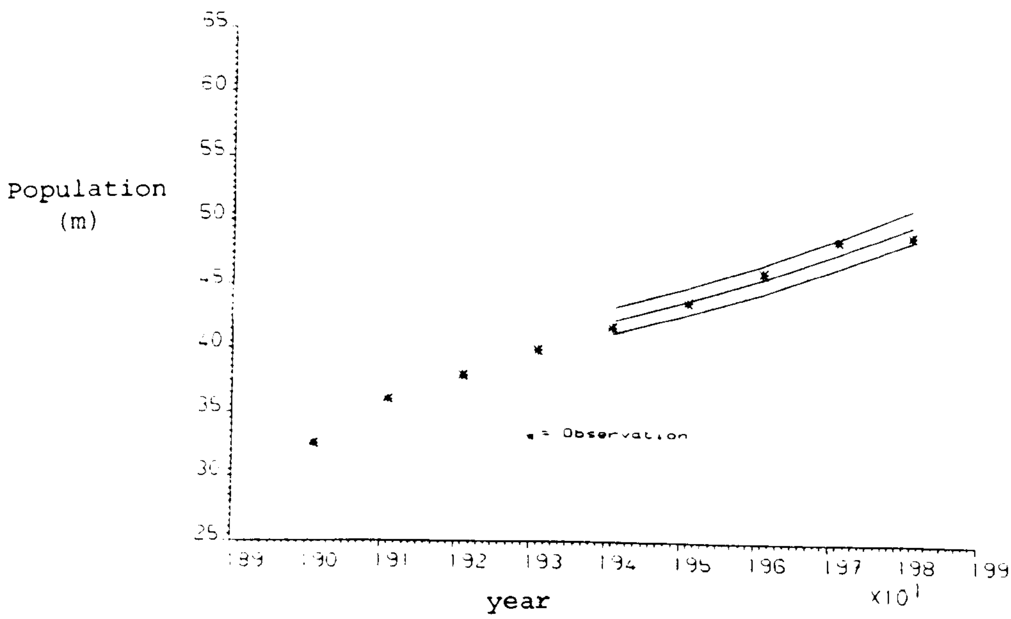
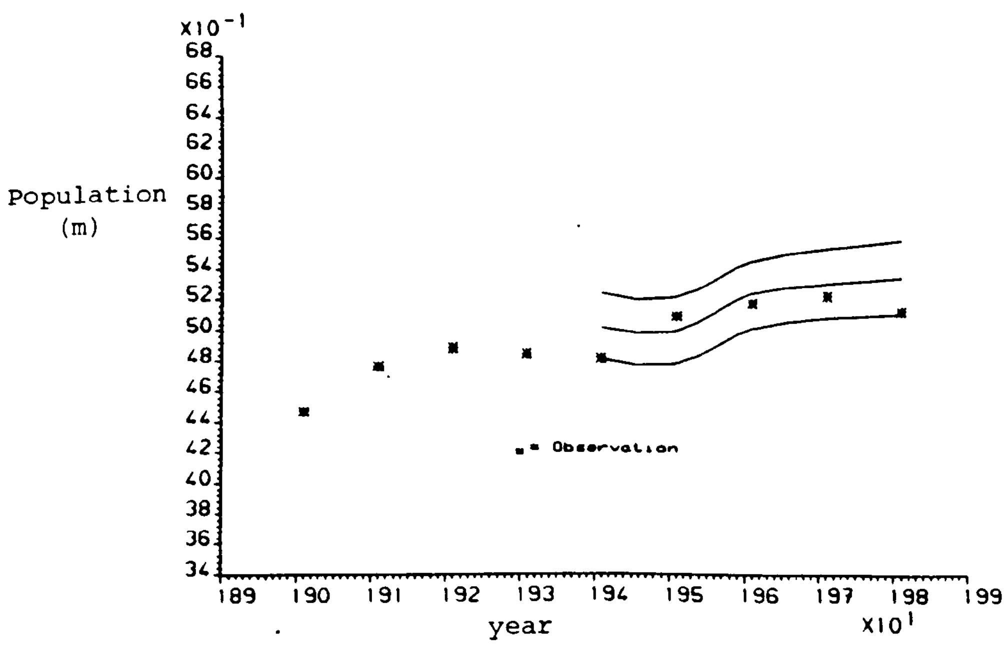


Figure 3.6.5: 1931 based forecasts of the Scottish population



dependent on the possibility of a change in the phase of growth occurring in that lead time, which would render future forecasts after that date unusable. To consider when this may occur is part of the art and judgment inherent in all forecasting. A qualitative framework by which this can be approached is developed in Chapter 6.

3.7 Comparison of logistic forecasts generated by different methods

Figure 3.7.1 illustrates, using boxplots, the absolute percentage errors in forecasts made by the various methods for generating forecasts based upon the logistic model for all forecasts made for all five populations since a base of 1930 up to a base of 1970. Figures 3.7.2, 3.7.3 and 3.7.4 compare the methods for forecasts with lead times of 10, 20 and 30 years. Figures 3.7.5, 3.7.6, 3.7.7 and 3.7.8 display the same information with the USA excluded. The boxplots visually illustrate the location and spread of the distribution of errors by vertical lines at the median and upper and lower quartiles. Non-overlapping notches (brackets) indicate that there is a significant difference between two boxplots. (For further details, see McNeil (1977).)

These Figures all show the extended Kalman filter method to give the most accurate forecasts (and significantly better in several cases).

Boxplots of Absolute Percentage Forecasts (USA included)

Figure 3.7.1: All forecasts

```
   ----
1 -I+)I---   ***000 0
   ----

    ---------
2 --I (+)   I----------  *** * *   *      0
    ---------

    -------------
3 --I  ( + )    I---------------    * **
    -------------

      -----------
4 ----I  ( + )  I-------------     * *
      -----------

  +---------+---------+---------+---------+
```

ONE HORIZONTAL SPACE = 0.70E 00
FIRST TICK AT 0.000

Key

1 = Kalman filter
2 = Asymptotic
3 = 4-parameter logistic
4 = 3-parameter logistic

Figure 3.7.2: 10 year ahead forecasts

```
     ---------
1 ---(  + ) I------        **     *
     ---------

    -----------
2 --I ,(  + )I-----------
    -----------

     --------------------
3 ---I   (    +     )   I--------------------
     --------------------

         -----------------
4  ------(   +    )      I----------------------  *
         -----------------

  +---------+---------+---------+---------+-------
```

ONE HORIZONTAL SPACE = 0.30E 00
FIRST TICK AT 0.000

Figure 3.7.3: 20 year ahead forecasts

```
    -----
1 --( +)I---            00   0
    -----

      -------------
2 ----I   (  +   )I---------                *
      -------------

       ------------------
3 -----I(     +    )    I-----------------------
       ------------------

        --------------
4 ------I     (   +  I)------------           * *
        --------------

  +---------+---------+---------+---------+-------
```

ONE HORIZONTAL SPACE = 0.40E 00
FIRST TICK AT 0.000

Figure 3.7.4: 30 year ahead forecasts

```
   ----
1 -I(+)---         0  00
   ----

        --------------------
2  -----I(      +       )  I---------------------
        --------------------

      ---------------------
3   --I   (       +      )I-----------------------
      ---------------------

        --------------
4   ----I   (    +   I)---              *       *
        --------------

  +---------+---------+---------+---------+-------
```

ONE HORIZONTAL SPACE = 0.50E 00
FIRST TICK AT 0.000

Boxplots of Absolute Percentage Errors (USA excluded)

Figure 3.7.5: All forecasts

```
              ----
   1 --(+)I----*
       ----

          ---------------
   2 ----I ( +  )        I----------------      *   *   *    *
          ---------------

         ---------------
   3 ---I    (  +  )   I-----------------
         ---------------

             -------------
   4 -------I    ( + )  I------------------
             -------------

    +---------+---------+---------+---------+---------+--

ONE HORIZONTAL SPACE = 0.40E 00
FIRST TICK AT        0.000
```

Figure 3.7.6: 10 year ahead forecasts

```
           --------
   1 ----( + )  I---------
           --------

          --------------
   2 --(I    +     )    I---------------
          --------------

            ---------------------------
   3  ---I  (          +         )     I-------------
            ---------------------------

                 ----------------------
   4  -------(-I       +        )       I---------------
                 ----------------------

    +---------+---------+---------+---------+-------

ONE HORIZONTAL SPACE = 0.20E 00
FIRST TICK AT       0.000
```

Figure 3.7.7: 20 year ahead forecasts

```
          -----
   1 ----(+ ) I------    *
          -----

               ------------------------
   2 ------(-I       +          )        I------------------
               ------------------------

              ----------------------------
   3 -------I     (           +          ) I---------------
              ----------------------------

                   ----------------------
   4 ------------I        (        +      I-)--------------
                   ----------------------

    +---------+---------+---------+---------+---------+-

ONE HORIZONTAL SPACE = 0.20E 00
FIRST TICK AT       0.000
```

Figure 3.7.8: 30 year ahead forecasts

```
        ----
   1 -(I+)I----
        ----

               ------------------------
   2   ----(-I         +          )     I---------
               ------------------------

             ----------------
   3    -(I      +       )   I-----------
             ----------------

               ---------------
   4    ----I    (      +     I-)-----
               ---------------

    +---------+---------+---------+---------+-

ONE HORIZONTAL SPACE = 0.40E 00
FIRST TICK AT       0.000
```

Key

1 = Kalman filter
2 = Asymptotic
3 = 4-parameter logistic
4 = 3-parameter logistic

3.8 Comparison of logistic based forecasts and "official" component based forecasts

Table 3.8.1 compares some "official" forecasts of Sweden, as published in the Swedish Yearbook (Statistisk Årsbok) by the National Central Bureau of Statistics, with those made by the logistic using the Kalman filtering procedure and with observed population.

By considering this Table the superior forecasting performance of the logistic is highlighted as forecasts from the logistic are only bettered (and only marginally) by the 1975 based component forecast. Similar results are found for the populations of England and Wales and Scotland (and hence Great Britain).

Forecasts of the future - 1990-2040 - made by the Kalman filtering procedure and made by the official agencies are illustrated in Figure 3.8.1.

As can be observed, except for Great Britain and England and Wales, there is little agreement between the component projections and the logistic forecasts. The logistic forecasts show high populations in the future for Sweden and Scotland, and lower forecasts for the USA. Which are the most likely can only be answered in time. A great deal depends on the continuation of sub-replacement fertility in Sweden and high emigration from Scotland for the logistic forecasts to be the poorest.

If it is considered that fertility will be around replacement level in the USA and allowing for ageing of the population, then it does seem that the logistic forecasts of the USA are on the low side. This illustrates that the model has not adequately accounted for escalation.

Table 3.8.1: Comparison of Kalman filter "logistic" forecasts with "official" projections* - Sweden

(i) 1955 based forecast

Year	Observed	Component projection	% difference	Logistic forecast	% difference
1960	7.498	7.388	-1.47	7.438	-0.76
1965	7.773	7.469	-4.07	7.699	-0.88
1970	8.081	7.546	-6.43	7.971	-1.31
M.A.P.E.			3.99		0.98

(ii) 1965 based forecast

Year	Observed	Component projection	% difference	Logistic forecast	% difference
1970	8.081	7.997	1.11	7.998	-0.98
1975	8.208	8.276	0.83	8.312	1.27
1980	8.318	8.531	2.56	8.444	1.51
M.A.P.E.			1.50		1.25

(iii) 1970 based forecast

Year	Observed	Component projection	% difference	Logistic forecast	% difference
1975	8.208	8.391	2.23	8.331	1.50
1980	8.318	8.703	4.63	8.465	1.77
M.A.P.E.			3.43		1.64

(iv) 1975 based forecast

Year	Observed	Component projection	% difference	Logistic forecast	% difference
1980	8.318	8.435	1.41	8.436	1.42

* Assuming that the central variant projection is the expected future.

Figure 3.8.1: Fifty year ahead forecasts compared to "Official Projections"

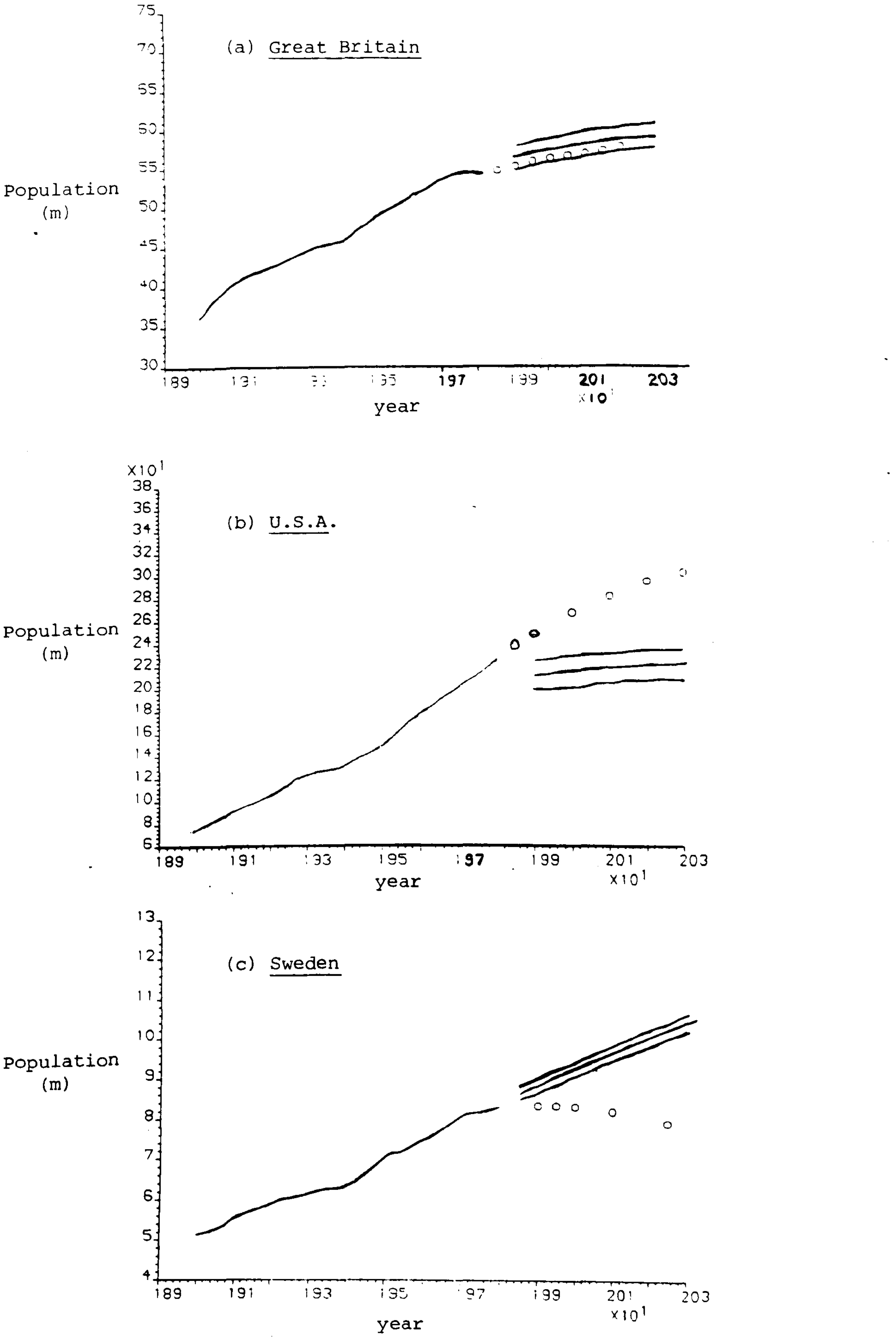
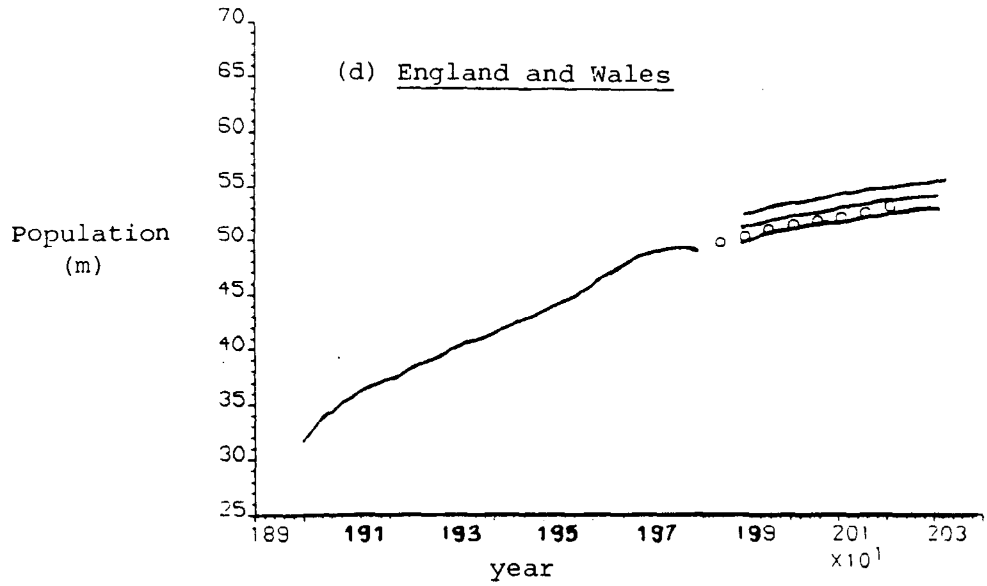
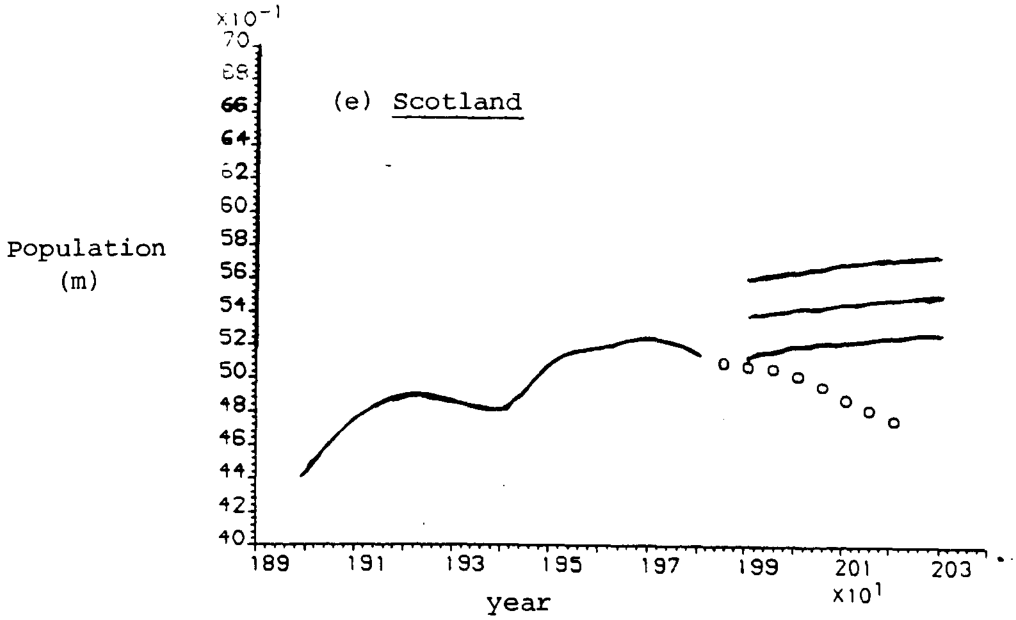


Figure 3.8.1 (continued)





3.9 Conclusion and Discussion

It would appear that, provided there are no major disruptions to the existing pattern of growth, forecasts produced from sigmoid models are satisfactory for the developed nations examined in this Chapter. These forecasts are best generated by using the extended Kalman filtering method which allows quick updating and generation of confidence limits.

Forecasts made in this way from a logistic model are as good (often significantly better) than any other method examined and compare favourably against the past performance of component projections (assuming that the real purpose of projections are to forecast).

Forecasts made by the logistic using the Kalman filter procedure were found to satisfy most of the criteria given by Meade (1984) in that logistic models have been shown to be statistically valid, have demonstrable forecasting ability and have forecast validity in terms of production of confidence intervals, while other means of forecasting sigmoids are found to fail on some of these grounds. The criteria of suitable saturation levels have not been considered in this Chapter as for human populations this is unlikely to be fixed and so little weight is given to the interpretation of the saturation level. This point will be returned to in Chapter 5.

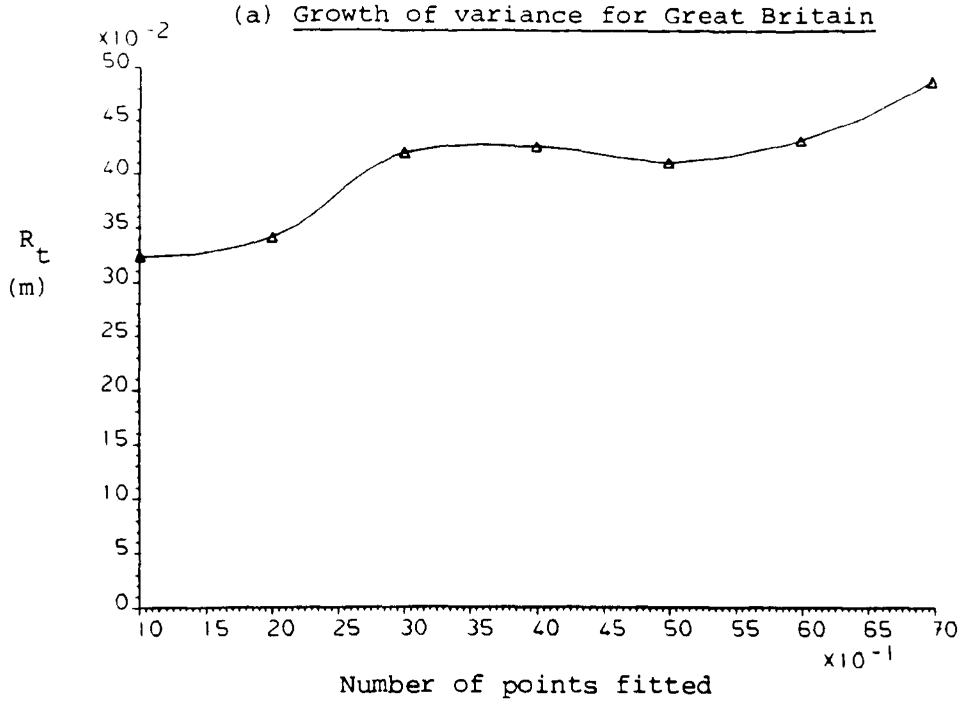
Having shown, in the previous Chapter, sigmoid type models to be suitable models of a population growth for a wide selection of nations, it is reasonable to expect that, by reformulating the sigmoids into recurrence expressions, the extended Kalman filter can be used to give forecasts of total population size for these nations.

There are still problems of giving useful forecasts if escalation is encountered as is demonstrated by the USA. However, the method of Kalman filtering allows the monitoring of the variance which can be used to detect when one has to be cautious in interpreting forecasts obtained. The use of monitoring the variance, R_t, is illustrated in Figure 3.9.1 where the variance at each iteration is plotted for the populations of Great Britain and the USA. If escalation is detected, then it is suggested that one should switch to the simpler forecasting scheme of the asymptotic method. The cases where escalation is required are best thought of as exceptions, and the need for escalation prompts a need for a more detailed study of the population.

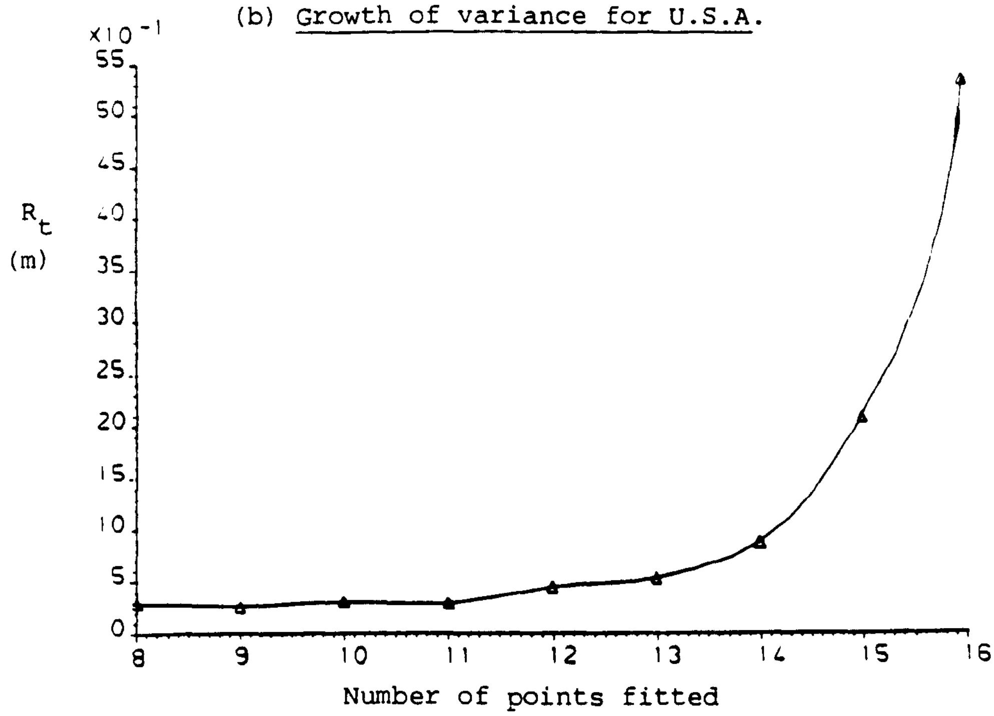
To produce forecasts for developing nations it would be best to use Murphy's asymptotic method, since it is considered likely that the future of population growth is likely to be volatile and there are problems of data shortages which preclude the use of least squares approaches.

Figure 3.9.1:

(a) Growth of variance for Great Britain



(b) Growth of variance for U.S.A.



Chapter 4 - Component-Population Modelling and Forecasting

4.1 Introduction

In this chapter an illustration of a "top down" modelling and forecasting methodology is presented. The essence of this is to derive estimates of the vital components of births, deaths and migration from the model fitted to total population.

In section 4.2 models of component-population of males and females are related to the total population. The problem of modelling and forecasting births is discussed at length in the section 4.3. An acceptable model is developed which is related to the total population model and it is shown that from this model forecasts with confidence limits can be obtained. The construction of this model of births also involves modelling life expectancy at birth which as is illustrated can be achieved using Sigmoids.

The death component and migration components are given brief consideration is section 4.4 and 4.5 and age specific fertility is briefly considered in the next section. In section 4.7 a methodology of forecasting age structure is presented and the method is illustrated for the population of England and Wales.

As one progresses through the discussion in this chapter it becomes apparent that the model of population on its own is insufficient to produce practical component population forecasts and better forecasts can be obtained by using the total population model in conjunction with models derived from past observations of the component-population. Hence a unified approach between "top down" and "bottom up" is adopted.

Throughout the chapter methods are presented with reference to "data rich" developed nations. However, discussion is made of appropriate procedures for population where less data is available.

In this chapter the aim is to illustrate one possible strategy and to highlight areas for future research.

4.2 Models of Male and Female Populations

The four parameter logistic model satisfactorily fitted the male and female populations of England and Wales over the period 1801 to 1981 (19 counts).

The parameters of the models are

	Males	Standard error	Females	Standard error
N	26.97	1.27	26.06	0.87
No	3.64	0.44	2.71	0.28
c	0.0206	0.0013	0.0239	0.0011
d	0.62	0.48	1.97	0.32
variance	2.7×10^{-4}		2.3×10^{-4}	

This gives an asymptotic ratio of females to males of 1.016.

Combining the models of the two sexes one forms the model for total population:-

$$p_t = \frac{53.03}{1 + \left(\frac{53.03}{6.35} - 1\right) e^{-0.0223t^*}} + 2.59 \; ; \quad t^* = \text{year} - 1801$$

The magnitude of the parameters P_∞, Po and d are slightly smaller than for the model fitted to the total population. However they are within the overlap of the standard errors. Thus one can model the population of the different sexes by logistic models.

Since addition of these models gives approximately the model for total population then dividing the parameters P_∞, Po and d of the model for total population by the sex ratio, will give approximately the model for each of the sexes.

4.3 Modelling and Forecasting Births

In this section models and forecasts of total births per year will be considered and related to models of total population that were developed in Chapter 2.

Fertility is the most important component of national population forecasting in the developed nations and the most volatile. Projections of fertility underpin cohort - component methods. Projections made by this method are very sensitive to judgements made to the likely fertility futures. Murphy (1984b) points out that often assumptions are made that fertility rates will become constant from about ten years ahead of the projection base and that this "asymptotic value has varied considerably over the period since 1955 in Great Britain leading to the projected number of births around the year 2000 varying from 600 thousand to 1.5 million". Such projections have obtained the reputation of unreliability if used as forecasts. Since empirical performance has been poor when projections are compared to observed births. The poor reputation is also due to intuitive reasons as often long term exponential growth is projected and projections have failed to reflect, that at least in the developed nations, fertility is by in large under conscious control. Ahlburg (1982) states that forecasts of live births are not more accurate than several naive methods and that performance of more recent forecasts is not better than earlier forecasts.

Fertility prediction is difficult because fertility depends on interactions and relative weights of inputs from economic, social, religious and political systems as well as attitudes to work and contraception.

Fertility can be measured in several ways, such as the crude birth rate or by taking account of age structure, age specific fertility rates, or the general fertility rate or possibly the most useful single figure the total fertility rate which is the average number of children that a woman would have if she experienced the fertility rates of a particular year throughout her reproductive span. In general demographers prefer rates which will allow standardised comparisons to measures given as absolute numbers per year.

However, Murphy (1980 and 1984a) has demonstrated that satisfactory fertility forecasts can be obtained if total births per year are used. Using this measure provides a natural link to the model of total population size.

The argument followed in this section is that given that total population follows a Sigmoid curve then births per year must have followed a similar curve, assuming that migration does not play a significant part.

The methodology developed is a time series one rather than a cohort approach that is used by many bottom up modellers, such as Easterlin (1968) or Keyfitz and Flieger (1968). Ermisch (1984) argues that the period approach is more suitable, particularly for nations which might be moving to an upper asymtote.

Lotka (1931) related the process of total population growth and the number of births per year through the renewal equation:-

$$P_t = \int_0^\infty B\ (t-a)\ P\ (a)\ da$$

Where $B(t-a)$ = Births at time t-a

a = age of person

$P(a)$ = probability of surviving to age a

P_t = Population size at time t.

From this Lotka (1931), obtained an expression for births at time t (B_t) as:-

$$B_t = \frac{P\infty}{Lo} P(t^*) + \frac{\lambda_2}{2!} P'(t^*) + \frac{\lambda_3}{3!} P''(t^*) + \frac{(\lambda_4 - 3\lambda_2^2)P^{iv}(t^*)}{4!} + \frac{(\lambda_5 - 9\lambda_2\lambda_3)}{5!} P^{v}(t^*) + \text{---------})$$

Where t^* = year-base year + λ_1; and λ_n is obtained from the moments of the lifetable ie.

$$\lambda_n = \frac{\int_0^\infty a^n\, P(a)da}{\int_0^\infty P(a)da} \approx \frac{\sum_{a=0}^{85} a^n\, P(a)}{\sum_{a=0}^{85} P(a)}$$

and P_t has a logistic form ie:-

$$P_t = \frac{P\infty}{1 + (\frac{P\infty}{Po} - 1)e^{-ct}} + d$$

Lotka shows that virtually all the births are given by the first element in the series. This Lotka refers to as the fundamental component and is given by: $\frac{P\infty}{e_o}(\tfrac{1}{2} + \tfrac{1}{2}\tanh \frac{ct^*}{2})$ which is equivalent to

$$\frac{P\infty}{e_o}\left(\frac{1}{1 + e^{-ct^*}}\right)$$

Where e_o ($\approx$ Lo) is life expectancy at birth. It is sensible to take female life expectancy at birth and to model and forecast only the female population, obtaining the total population by multiplying by one plus the sex ratio.

The only additional knowledge required to model the bulk of births is the first two moments of the life table. Keyfitz (1968) confirms this and writes Lotka's model more succinctly as

$$B_t = \frac{P(t+\bar{\mu})}{e_o} - \tfrac{1}{2}\bar{\sigma}^2 \frac{P''(t+\bar{\mu})}{e_o}$$

where $\bar{\mu}$ = the "mean age of the living" in the life table

$$= \frac{L_1}{L_0} \text{ and } \bar{\sigma}^2 = \left[\frac{L_2}{L_0} - \left(\frac{L_1}{L_0}\right)^2\right] \; ; \; L_n = \int_0^\infty a^n P(a)\,da$$

Keyfitz finds this model to be robust to changes in the parameter values.

To implement this procedure over a protracted period, changes in the moments with time should be reflected in the modelling.

Taking first life expectancy at birth, the graph of English and Welsh life expectancy over time is displayed in figure 4.3.1. This shows that throughout the previous century there has been a continued improvement, at first quickly and in more recent years at a slower rate.

Changes in life expectancy over years at birth can be described by a Sigmoid. Hart and Hertz (1944) uses e_o as an index of the social progress of a nation and found that this can be described well by four parameter logistic models, which in turn can be used to give forecasts of future e_o.

Sigmoids were fitted to e_o as recorded in lifetables for different calendar years. Since e_o is going to be modelled as a function of time it is given the subscript t. The four parameter logistic was found to be the most suitable model ie:-

$$e_{ot} = \frac{E_\infty}{1 + (\frac{E_\infty}{E_o} - 1)e^{-ct^*}} + d$$

A summary of the models obtained for the population of England and Wales, France, Japan, Scotland, Sweden and the USA is given in table 4.3.1.

Figure 4.3.2 illustrates how the magnitude of the upper asymptote (E_∞ +d) has altered for England and Wales as new information is obtained. Figure 4.3.3 shows the case for Swedish e_{ot}.

There is quite a bit of variability in the fits of the logistic models to e_{ot} and a case can be made for fitting escalated curves which might coincide with medical advances, such as smallpox vaccination programmes or in preventative medical care such as anti-smoking campaigns in Great

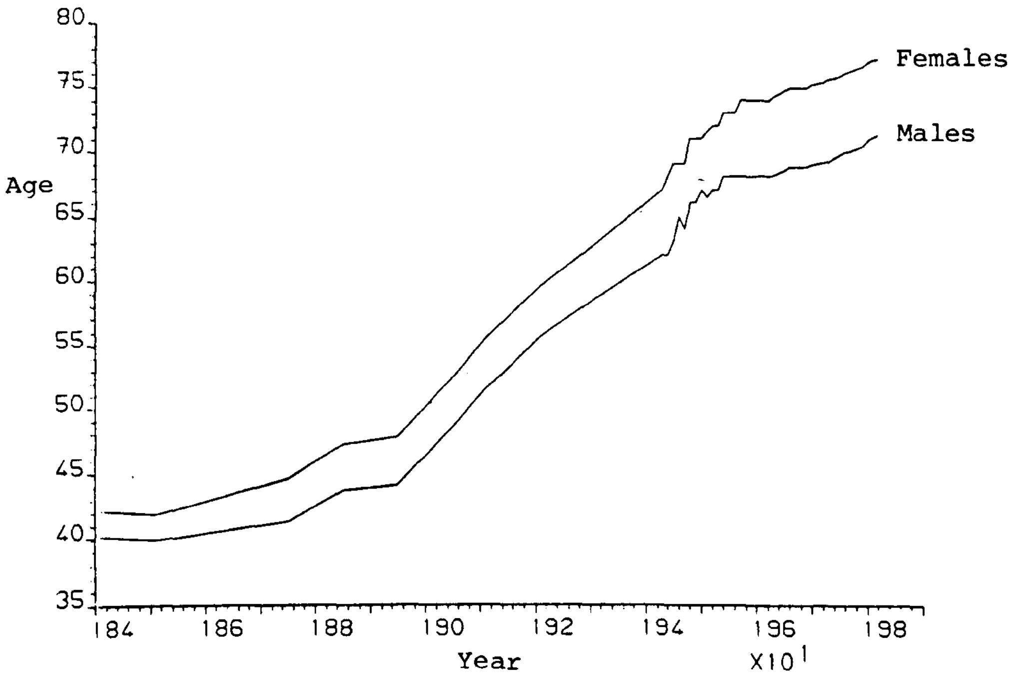


Figure 4.3.1 - English and Welsh Life Expectancy at Birth

Figure 4.3.2 - Variations in E∞+d for England and Wales

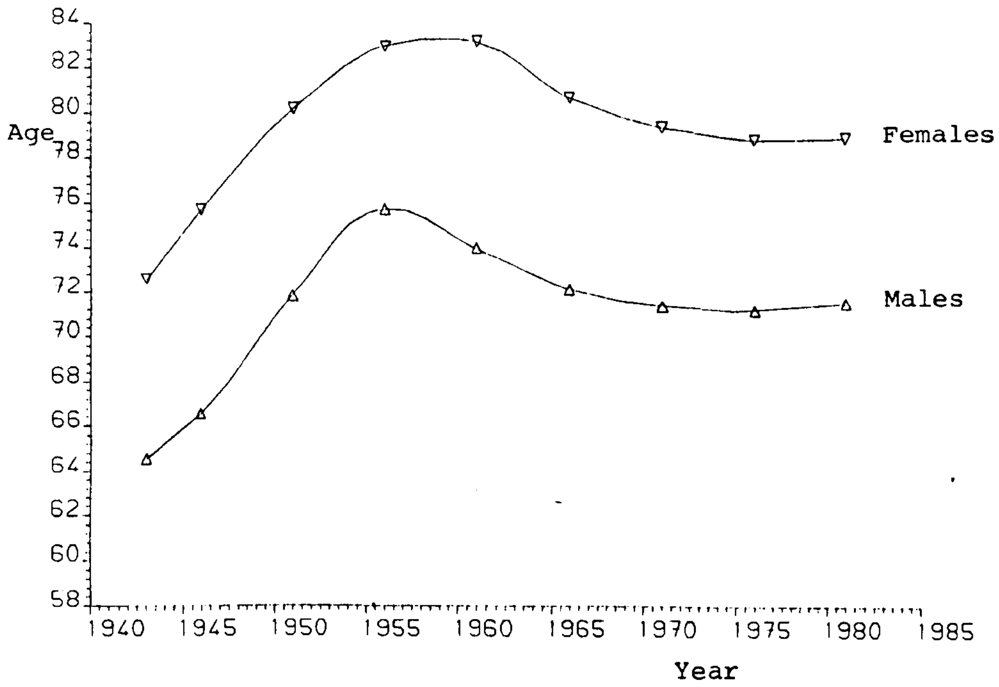
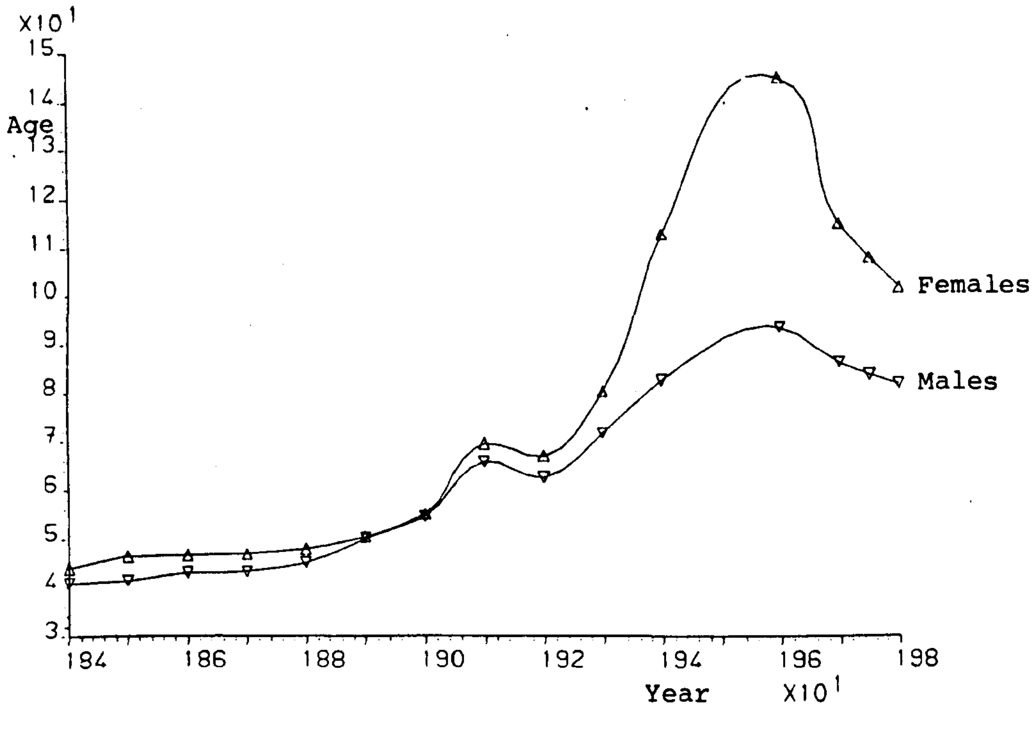


Figure 4.3.3 - Variations in E∞+d for Sweden



Britain in the 1970's. The effect of such escalations were found to be slight when compared to one model fitted to the whole series. As a result this complexity is not pursued.

From the models listed in table 4.3.1 forecasts can be made.

At this point one can carry out a little exploratory investigation as to the usefulness of the model. Logistic models are fitted to the birth series using the non-linear regression algorithm discussed in Chapter 2 (section 2.4). The model can be confirmed for asymptotic cases in that the upper asymptote of the model derived from the fit should be approximately equal to $P\infty/E\infty$.

For the English and Welsh birth series 1838-1981 the model arrived at is

$$B_t = \frac{262.5}{1 + (\frac{262.5}{270.0} - 1)e^{-0.0002t^*}} + 430.0$$

Where B_t = births in thousands and t^* = year - 1801. This gives an upper asymptote of 0.692 million births per year.

The asymptote of the population model divided by the asymptote of e_{ot} $= \frac{55.5}{78.95} = 0.703$m. This gives the percentage discrepancy of the B_t model of only 1.6%. Thus it seems to be worthwhile exploring this model further.

Table 4.3.1 - Logistic models of Life Expectancy

Population	Female Life Expectancy Parameters		Male Life Expectancy Parameters
England and Wales	E_∞	56.940	43.252
1841-1981	Eo	2.979	0.954
t* = year-1801	c	0.0255	0.032
number of	d	29.850	33.858
observations = 47	variance	$8.3\text{x}10^{-4}$	$1.5\text{x}10^{-3}$
France	E_∞	48.748	38.498
1851-1980	Eo	0.431	0.1393
t* = year-1801	c	0.036	0.0420
number of	d	36.441	36.863
observations = 25	Variance	$5.1\text{x}10^{-4}$	$6.2 \text{ x } 10^{-4}$
Japan	E_∞	36.758	32.521
1895-1983	Eo	0.000001	0.000001
t* = year-1801	c	0.112	0.119
number of	d	43.624	42.437
observations = 12	variance	$1.0\text{x}10^{-3}$	$1.2\text{x}10^{-3}$
Scotland	E	38.070	36.098
1861-1981	Eo	0.144	0.301
t* = year-1981	c	0.0424	0.037
number of	d	41.903	37.744
observations = 12	variance	$1.0 \text{ x } 10^{-3}$	$2 \text{ x } 10^{-3}$
Sweden	E_∞	68.000	49.768
1861-1981	Eo	1.786	0.748
t* = year-1801	c	0.0191	0.0257
number of	d	33.405	31.686
observations = 24	variance	$2.4\text{x}10^{-4}$	$3.3\text{x}10^{-4}$
USA	E_∞	68.473	66.638
1914-1979	Eo	2.226	9.208
t* = year-1790	c	0.0295	0.0189
number of	d	15.767	12.242
observations = 61	variance	$4.9\text{x}10^{-4}$	$6 \text{ x } 10^{-4}$

Other life table moments can be modelled in a similar manner. For example changes in the moments of L_{1t} and L_{2t} for England and Wales over the period 1846 to 1981 were modelled by the models:-

$$\hat{L}_{1t} = \frac{1706.3}{1+(\frac{1706.3}{5.8}-1)e^{-0.04t^*}} + 1723.9$$

$$\hat{L}_{2t} = \frac{107656.4}{1+(\frac{107656.4}{287.2}-1)e^{-0.046t^*}} + 84351.8$$

$$t^* = \text{year} - 1801$$

Observed values of L_{1t} and L_{2t} were obtained by using the Reed-Merril transformation (see Shryock et al (1976), for details) to convert five yearly grouped age specific mortality at time $({}_5M_{at})$ into the conditional probability of dying by a certain age at time t $({}_5q_{at})$, given the person was alive at t-5 which in turn is converted into the probability of surviving to age at time t $(P_t(a))$. From this L_{nt} can be approximated ie

$$L_{nt} = \sum_{a=o}^{85} a^n P_t(a)$$

This involved the following procedure:

$${}_nq_{at} = 1-e^{-n\,{}_nM_{at} - \alpha n^3\,{}_nM^3_{at}} \longleftarrow \text{Reed-Merril Transformation}$$

where n = size of age group (in this case n = 5)

α is thought best to be 0.008

t = calendar year for which ${}_nM_a$ applies.

This requires slightly different treatment for younger ages ie:

$$\text{for } {}_1q_{ot} = 1 - e^{-1\,M_{ot}\,(0.9539 - 0.5509\,{}_1M_{ot})}$$

$$\text{and for } {}_4q_{1t} = 1-e^{-4\,{}_4M_{1t}\,(0.9806 - 2.709\,{}_4M_{1t})}$$

Then $P_t(o)$ is set to one and $P_t(a)$ is found from:-

$$P(a)_t = (1-{}_nq_{at})\,P_t(a-n)$$

Thus for ages >5

$$L_{1t} = \underline{a}^T{}_t\,P_t(a) \text{ and } L_{2t} = [{}_n\underline{a}^2{}_t]^T\,P_t\,(a)$$

Where $\underline{a}$ is an 'age' vector ie:-

$$\underline{a}^T = 5, 10, 15, 20 - - - - - 75, 80, 85$$

These moments and the logistic model of the English and Welsh population, 1801 to 1981 (see Chapter 2) were then input into the expression:-

$$B_t = \frac{P_{(t+\bar{\mu}t)} - \frac{1}{2}\,\bar{\sigma}_t^{\,2}\, P''_{(t+\bar{\mu}t)}}{e_{ot}}$$

to obtain estimates of English and Welsh births over the period 1841 to 1981. These are compared to actual births in figure 4.3.4.

The performance of the elaborate model developed is rather poor. It follows the general trend, but fails to reflect the magnitude of the major swings in the birth series. This is unfortunate especially since as Ryder (1969) stated that "the future of fertility is likely to be increasingly bound up with questions of fluctuation rather than of trend" (referring to post-transition populations). Commenting on populations which display fluctuations in births (all populations discussed in this chapter do) Lee (1974) gives further condemnation of such an approach as outlined above by stating that "the ability of demographic analysis to detect subtle trends is not useful for population predictions". In practice this may be an unjust criticism since from a users view point a notion of the future trend of births is likely to be sufficient provided that it is accompanied with indications as to the degree of uncertainty as to how far the future births might be expected to deviate from this trend. This shows up the main fault in the above approach in that it does not seem possible to generate confidence limits for such a many parameterised non-linear model. Thus for forecasting alternative methods need to be examined.

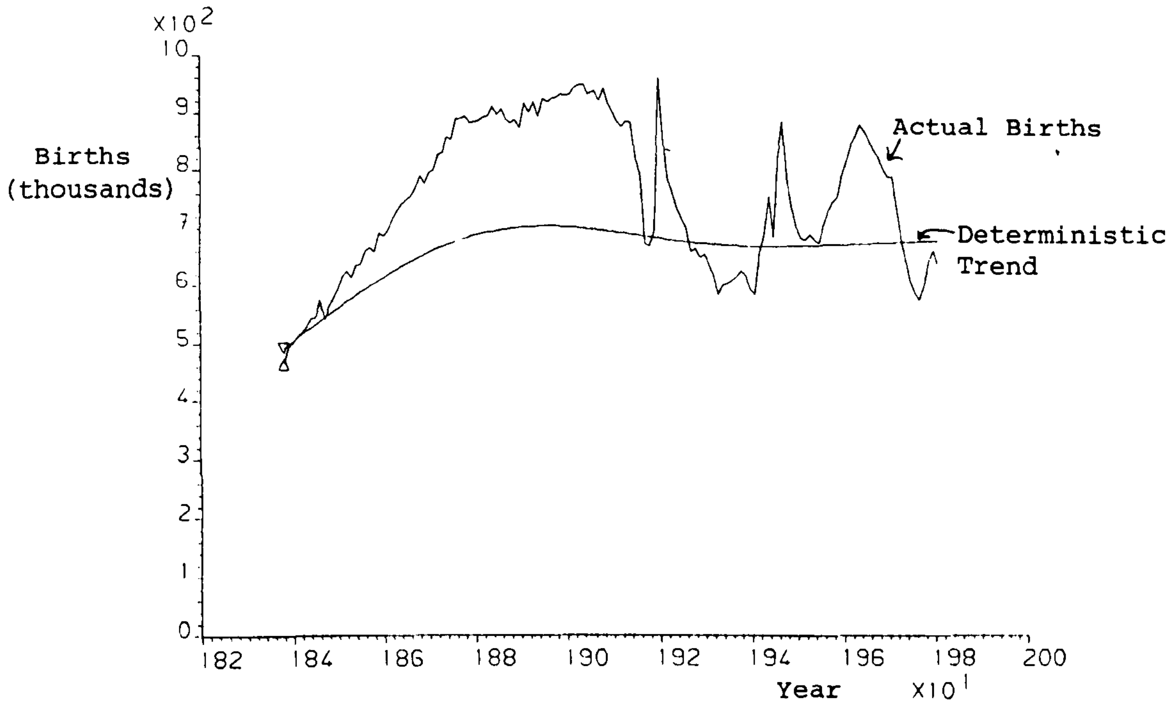


Figure 4.3.4 - English and Welsh Births

ARIMA models (see chapter 3, section 3.4) are particularly interesting for modelling birth time series as they make explicit use of the inherent autocorrelation structure of the series, which is an advantage over many other methods since as Lee(1974) states "to ignore the autocovariance structure of fertility is to throw out the most useful available information". Other perceived advantages of ARIMA models are, by virtue of how the model is identified, in allowing the data to "speak for itself" and the "automatic" generation of confidence limits.

For these reasons and increased availability of computer implementations of ARIMA models, such models are gaining popularity in demography for modelling vital events, for which at least in the developed nations there is sufficient data to support such procedures. However, when Lee (1974) attempted to model USA births (1917 to 1972) he found the results to be disappointing and it was problematic in trying to achieve stationarity in the time series. Therefore Lee concluded "the severe nonstationarity implied by the necessity of second differencing makes this approach unattractive".

Saboia (1977) on the other hand found ARIMA models to be good for modelling and forecasting the birth series of Norway and attributes demographic significance to the largest roots of the ARIMA models relating them to the eigenvalues of the Leslie matrix (see Keyfitz 1968).

The first eigenvalue is associated with the eventual rate of increase of the female population and the next two are related to generation length. To incorporate these parameters Saboia is forced to consider non-parsimonious models, (since five ARIMA coefficients are required).

McDonald (1980) is critical of Saboia's approach stating that "the non-parsimonious nature of Saboia's models raise a serious methodological problem. The difficulty is that if a nonparsimonious model is selected then it is always possible to find an ARIMA representation with the required characteristics whatever the properties of the time series analysed." McDonald (1980) shows that Saboia's models are not parsimonious containing common factors and insignificant coefficients and so the models cannot be regarded as evidence in favour of Saboias argument. But McDonald (1979) reporting on project FORECAST which uses ARIMA models to forecast Australian birth favours the use of such models for long term forecasting.

Murphy (1984b) explored the use of ARIMA models and found that ARIMA forecasts performed much more satisfactorily than other methods both in the accuracy of forecasts and that the confidence intervals are broadly in agreement with the variability found in the series. He states that ARIMA models "perform better than forecasts made using separate information on rates and populations at risk and that births (annual counts) form a better basis for forecasts of births than alternative measures."

Murphy (1984b) modelled the fertility series of England and Wales, USA, France and Norway and in all cases found ARIMA models to be best for short periods and to be good for long periods.

Thus there does appear to be merit in investigating the use of ARIMA models further. Other applications of ARIMA models to vital demographic events are commented on by Long and McMillen (1984) (models of US births) Land and Cantor (1983) (who model seasonal effects on birth and death rates) and by de Beer (1983) (who examine ARIMA type models for changing cohort data over time).

Applying ARIMA models to the birth series confirms that useful forecasts can be generated for example table 4.3.2 displays forecasts made from an ARIMA (1,1,0) model fitted the Swedish crude birth rates 1926-1975 and compared to observed births.

Before investigating ARIMA models further it is important to stress that there is no completely objective way of making forecasts and ARIMA models are especially vulnerable to forecaster bias in model identification and treatment of outliers. (Pittenger 1978, discusses this further). The problem of chosing the optimal ARIMA model specification from the verity of possible models given in the identification phase is not so important for short term forecasts, as these tend to be similar, but long term forecasts from the different models will often tend to be very different.

Table 4.3.2 - Forecasts of Swedish Crude Birth Rates

Model fitted over the period 1926 to 1975

ARIMA (1,1,0)

$$= \underset{(0.156)}{(1 + 0.554B)}(1-B)Z_t = e_t$$

Forecasts

Year	Forecasts	Observed	% error
1976	12.50	12.14	2.97
1977	11.97	11.80	1.44
1978	11.81	11.40	3.60
1979	11.67	11.71	-0.34
1980	11.66	11.818	-1.35
		M.A.P.E.	1.94

Thus one can write the model for births as

$$B_t = DB_t + R_t ;$$

where DB_t is the deterministic trend and R_t is the stochastic residual part, perhaps reflecting exogenous changes in society, technology and the economy.

In the cases above differencing or transformations were required to achieve stationarity. In such series Lee (1979) suggested an alternative strategy that a "logistic trend" be removed from the series and the residuals modelled by ARIMA techniques. This would suit a "top down" approach.

This modelling was applied to the English and Welsh birth series (1838 to 1981). Initially the birth series was modelled up to 1921 and modelling repeated by extending the time period up to the next census until 1981 is reached. For each iteration forecasts were made to 1981 and compared to later observations.

The stochastic ARIMA models can be obtained by using the MINITAB statistical package, or in this case by NAg10 library routines incorporated into a FORTRAN IV programme BIRTH which is listed in Raeside (1985). This programme automatically generates 50 and 90% confidence limits around forecasts.

The deterministic part of the models are displayed in table 4.3.3a and the stochastic part is laid out in table 4.3.3b. Examination of the residuals, and their auto correlation structure from these models shows that stationarity is approximated, although around the war years there were some outliers which should perhaps be removed (Murphy (1984) applies some ad hoc treatment of outliers in his modelling of this series). However, it was found that treatment for outliers by redistributing the births around this period made very little difference and so no treatment was employed.

These models were found to give satisfactory forecasts as most future observed births (except for outliers associated with the baby boom after the second world war) lie within the 90% confidence intervals and the expected path follows the direction of the trend of births. Figure

4.3.5 displays the model, observed births and fifty year forecasts for models fitted 1838 to 1931, 1838 to 1971 and 1838 to 1981.

A striking feature of the plots exhibited in figure 4.3.5 is that the confidence intervals are wide (±30% for 1931 based forecasts).

Table 4.3.3a - The Form and Parameters of the Deterministic Equation for England and Wales

Using the form stated by Keyfitz (1977) pp 213:-

$$DB_t = \frac{P(t+\bar{\mu}) - \tfrac{1}{2}\bar{\sigma}^2\, p''(t+\bar{\mu})}{e_{ot}} \quad \text{as an approximation}$$

The parameters of p_t and e_{ot} for each of the periods are as laid out below:-

Period	p_t: p	p_o	c	d	e_{ot}: E_∞	E_o	γ	ψ
1838-1921	79.19	8.88	0.0169	0.0	46.47	0.11	0.0468	41.21
1838-1931	58.92	6.96	0.0205	1.96	30.51	0.05	0.0564	41.55
1838-1941	52.50	6.08	0.0226	2.87	31.06	0.05	0.0556	41.52
1838-1951	50.56	5.76	0.0234	3.20	40.24	0.16	0.0442	40.86
1838-1961	50.98	5.84	0.0232	3.12	41.50	0.18	0.0431	40.785
1838-1971	50.98	6.24	0.0223	2.70	38.02	0.11	0.0480	41.21
1838-1981	52.77	6.19	0.0224	2.75	56.94	2.974	0.0255	29.85

p_t and e_t are both logistic models with upper asymptotes $p_\infty + d$ and $E_\infty + \psi$ respectively

The First Moment - over the years 1846 to 1981 'best' model is:-

$$\hat{L}_{1t} = \frac{1706.3}{1+\left(\frac{1706.3}{5.8}-1\right)e^{-0.04\,t^*}} + 1723.9$$

The second Moment - over the years 1846 to 1981 'best' model is:-

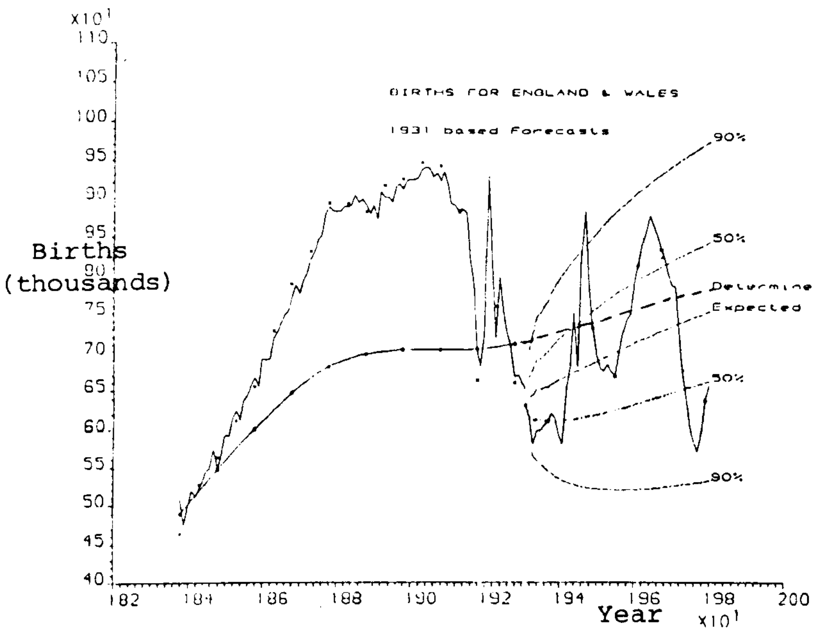
$$\hat{L}_{2t} = \frac{107656.4}{1+\left(\frac{107656.4}{287.2}-1\right)e^{-0.046t^*}} + 84351.8$$

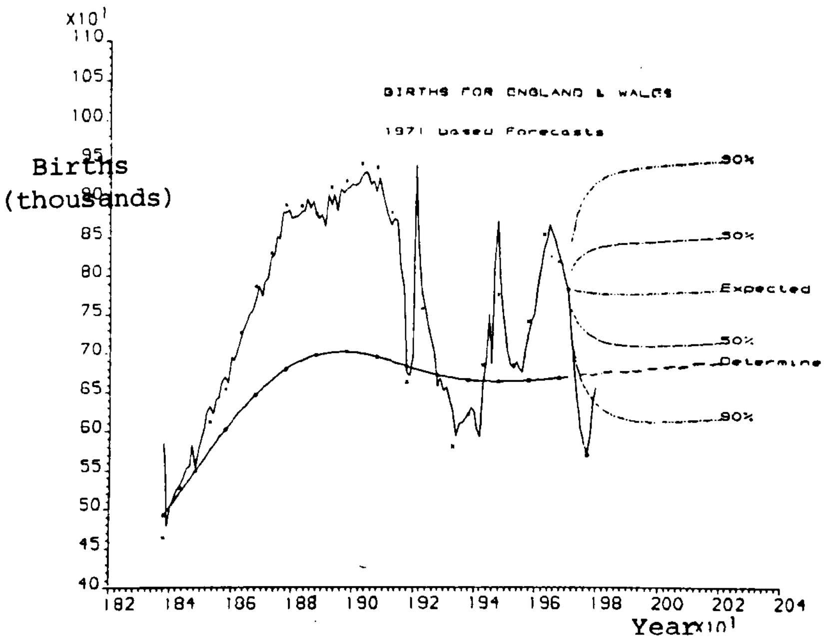
t^* = year - 1801

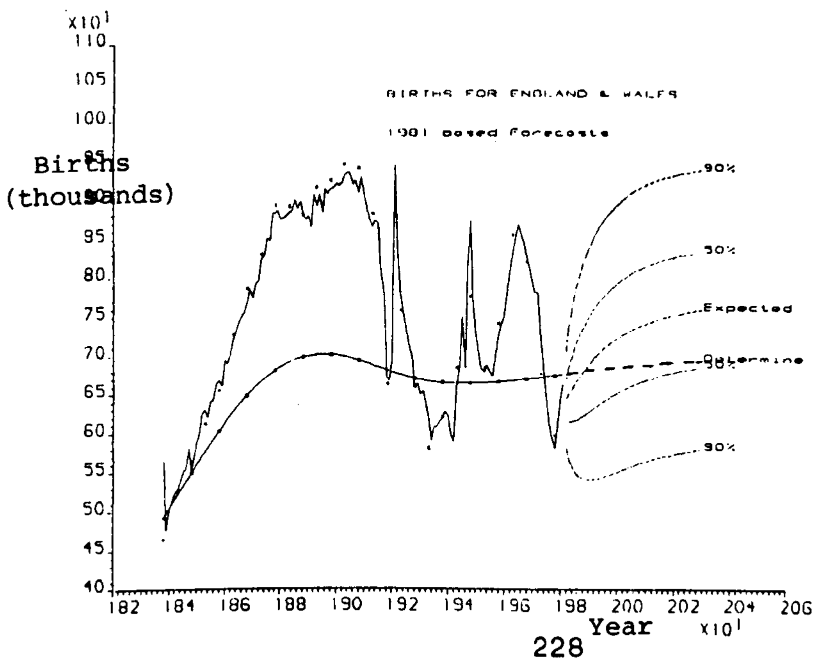
Table 4.3.3b - The Arima Models for the Residual Component of English and Welsh Births (standard errors are in brackets)

PERIOD	ORDER OF MODELS	PARAMETERS	SUM OF SQUARES	COMMENTS	EVENTUAL FORECAST FUNCTION
1838-1921	(4,0,0)	$Z_t=0.872Z_{t-1} + 0.187Z_{t-2} - 0.909Z_{t-3}$ (0.112) (0.216) (0.235) $+ 0.823Z_{t-4} + e_t$ (0.146)	83418	AR(2) term is not significant. Some non-stationary problems	$F_t=0.87F_{t-1} + 0.19F_{t-2} - 0.91F_{t-3} + 0.82F_{t-4}$
1838-1931	(4,0,0)	$Z_t=0.870Z_{t-1} + 0.032Z_{t-2} - 0.219Z_{t-3}$ (0.101) (0.135) (0.135) $0.296Z_{t-4} + e_t$ (0.102)	113913	AR2 and 3 terms are not significant. ARIMA (2,1,2) is also good	$F_t=0.87F_{t-1} + 0.03F_{t-2} - 0.22F_{t-3} + 0.30F_{t-4}$
1838-1941	(4,0,0)	$Z^t=0.876Z_{t-1} + 0.025Z_{t-2} - 0.220Z_{t-3}$ (0.096) (0.128) (0.128) $+ 0.301Z_{t-4} + e_t$ (0.096)	115951	AR2 and 3 not significant ARIMA (2,1,2) is also good	$F_t=0.88F_{t-1} + 0.03F_{t-2} - 0.22F_{t-3} + 0.30F_{t-4}$
1838-1951	(4,0,0)	$Z_t = 0.881Z_{t-1} - 0.003Z_{t-2} - 0.113Z_{t-3}$ (0.094) (0.126) (0.126) $+ 0.191Z_{t-4} + e_t$ (0.095)	171661	AR2 and 3 not significant ARIMA (2,1,2) is also good	$F_t= 0.88F_{t-1} - 0.01F_{t-2} - 0.11F_{t-3} + 0.194F_{t-4}$
1838-1961	(4,0,0)	$Z_t = 0.906Z_{t-1} - 0.008Z_{t-2} - 0.109Z_{t-3}$ (0.090) (0.122) (0.122) $+ 0.172Z_{t-4} + e_t$ (0.091)	178606	AR2 and 3 terms not significant, AR(1) would be good - but not not so good at forecasting	$F_t =0.91F_{t-1} - 0.01F_{t-2} - 0.11F_{t-3} + 0.16F_{t-4}$
1838-1971	(1,0,0)	$Z^t = 0.928Z^{t-1} + 94.11$ (0.034) (41.34)	185670	-	$F_t =0.93F_{t-1} + 94$
1838-1981	(1,0,0)	$Z_t = 0.953Z_{t-1} + e_t$ (0.037)	126667	AR(4) is also a good model	$F_t = 0.93F_{t-1}$

Figure 4.3.5 - English and Welsh Births







These would be narrower if there had been treatment for outliers as the variance would be reduced (but doing this had little practical effect). The width of the intervals highlights the inherent variability and lack of precision that can be obtained in forecasting births and so has the advantage of making this clear to decision makers and to indicate that a high degree of flexibility must be incorporated into their decision frameworks. (McDonald (1979) discusses this point further).

As can be seen from the illustration of the 1971 forecasts in figure 4.3.5 the forecasting performance is poor as observed births over the period 1972 to 1981 "fall" out of the bottom of the 90% confidence interval. This possibly indicates that fertility was temporarily "reduced" and is well below the expected path. If this trend continued then one should consider that the model is predicting too high values and investigate the possibilities of de-escalation. However, the observed births re-enter the anticipated region and the low departure might be seen as "balancing" the previous high and since there is no evidence of de-escalation from the model fitting the conclusion reached is that future English and Welsh births will rise and move into the anticiapted region. By comparing the forecasts 1972-1981 for the ARIMA plus deterministic model to observed births the MAPE is disappointingly high at 23.77%. However, this is an improvement over the official projections of births (OPCS,1972) which have a MAPE of 27.53%. (Note that the performance of the deterministic extrapolation is much better having a MAPE of 8.40%).

Thus although the forecasts are poor the method based on the logistic model is superior to official cohort component methods. This is confirmed when the logistic based model is updated by inputs of births over the period 1972 - 1975. This changed the ARMA part of the model for the period 1838 to 1975 to:

$$\text{ARMA } (1,0) \qquad \underset{}{Z_t = \underset{(0.0295)}{0.9504}Z_{t-1} + e_t}$$

(The deterministic part remaining unchanged).

The percentage error of this model and of OPCS central variant projections made in 1975 (CPRS 1977) are displayed in table 4.3.4. From this table one observes that the logistic - ARMA model is better overall than the OPCS projections. However, the OPCS projections are better for very short lead times.

Table 4.3.4 - Logistic ARMA forecasts Versus OPCS 1975 based projections

Year	Percentage error in forecasts/projections	
	Logistic - ARMA	OPCS
1976	3.95504	-0.0513
1977	7.4034	1.1769
1978	3.1555	-2.9175
1979	- 3.0219	-6.5831
1980	- 5.2307	-4.1597
1981	- 1.4987	5.1222
1982	0.3030	12.4781
1983	0.2074	17.3105
	MAPE 3.10%	6.22%

1981 based forecasts (see figure 4.3.5) indicate that the number of births per year is expected to rise.

In the very long term one would speculate that movement shall be close to replacement level which is given by the deterministic level. The validity of such speculation is considered in relation to current demographic theory in Chapter 5.

In the above modelling the variance term ($\bar{\sigma}^2$) is very small implying that the term involving the second derivative of the population model adds little to the model (less than 3%). Hence, it is suggested that this term be dropped and incorporated into the ARIMA model. Murphy (1980 and 1984a) simplifies the approach even further by approximating $\bar{\mu}$ by $\frac{1}{2}e_{ot}$ and shows the method to be relatively insensititve to errors. Without the ARIMA part Murphy shows that the model:-

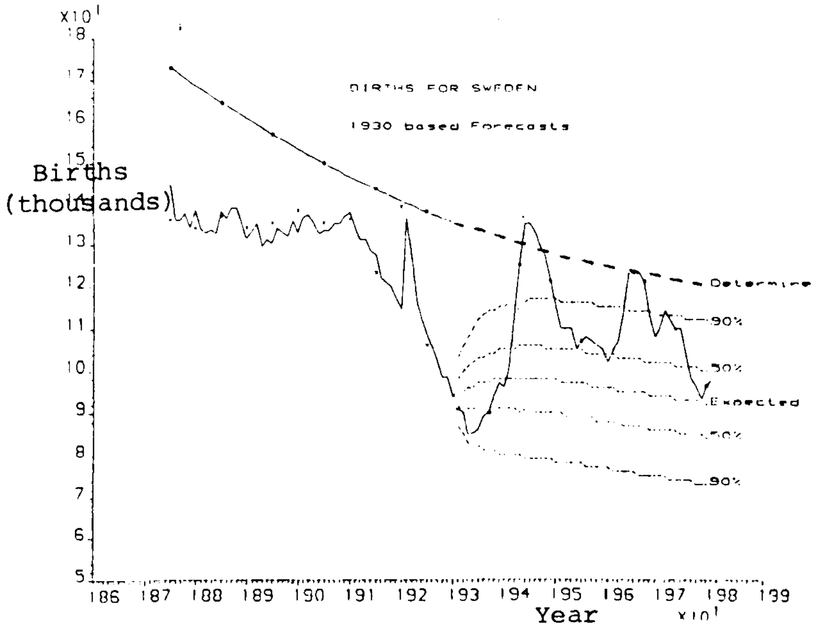
$$B_t = \frac{P(t+\frac{1}{2}e_{ot})}{e_{ot}}$$

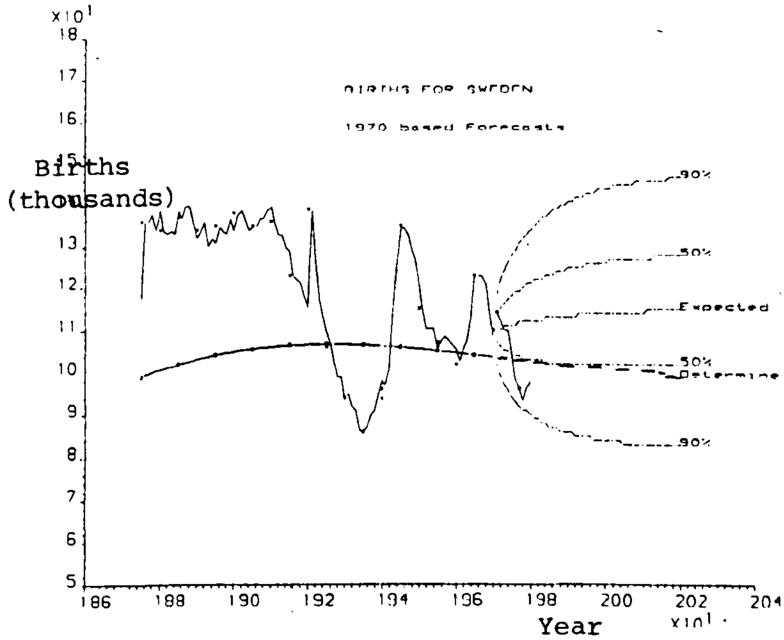
has reasonable forecasting performance for British births and that the root mean square error of forecasts is one-third greater for continuing the present value and two-thirds greater for official projections. This greatly simplifies the calculation of the deterministic part.

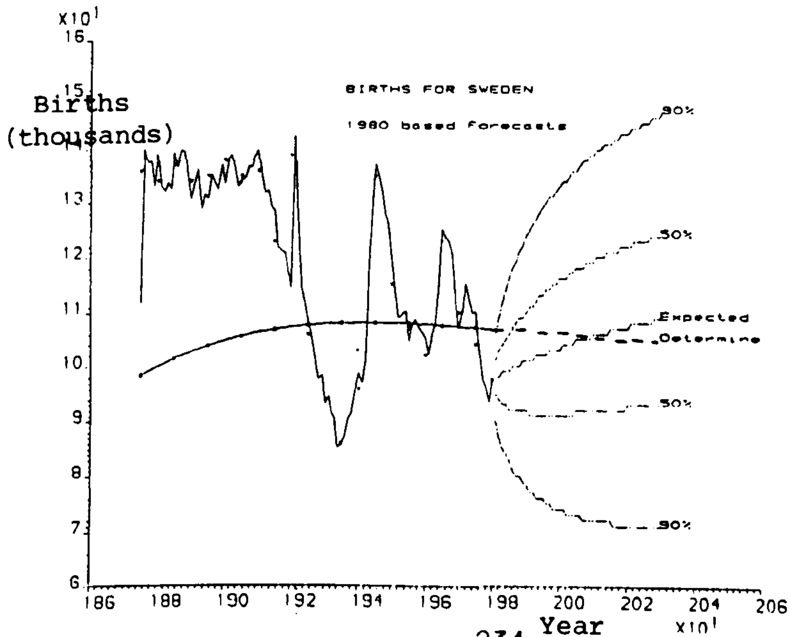
Thus for other models of births the above simple model is used for the deterministic trend leaving the ARIMA part to "soak" up the error.

This is applied to the birth series of Sweden, USA, Scotland, France and Japan. Forecasts made for Sweden are displayed in figure 4.3.6 for bases in 1930, 1970 and 1980. Figure 4.3.6 showing the 1930 based forecasts indicates that the deterministic part predicts far too high values of births, however the ARIMA part has successfully adapted to adjustments downwards. The bulk of future births are within the 90% confidence intervals, apart from births immediately after 1945. which were high possibly as a result of delaying births during the second world war and should be perhaps treated as an outlier. Births also appear unexpectedly high in the mid 1960s which may suggest escalation of the model which was also indicated by the exploratory analysis carried out prior to fitting the logistic model to population (see chapter 2 section 2.7). Since the upward departure is short lived escalated models make little difference, ratifying the use of one continuous model. For the 1970 based forecasts the expected path is close to the observed births for the first 4 years and thereafter predicts rising births while the observed trend appears to be falling. The deterministic path seems to give better long term predictions. As can be observed from figure 4.3.6 the expected fifty year future for births is to rise to slightly above the deterministic path which would put Sweden on course for long term replacement level fertility.

Figure 4.3.6 - Swedish Births

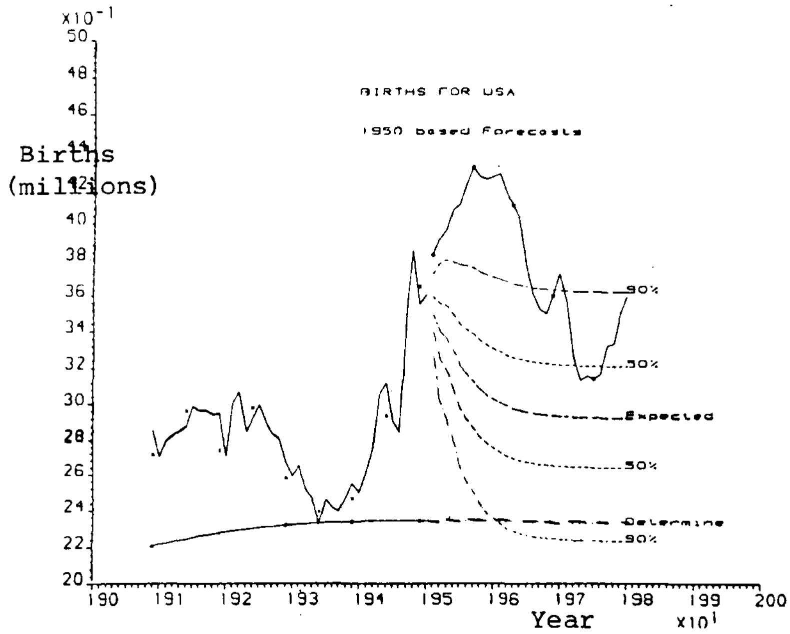


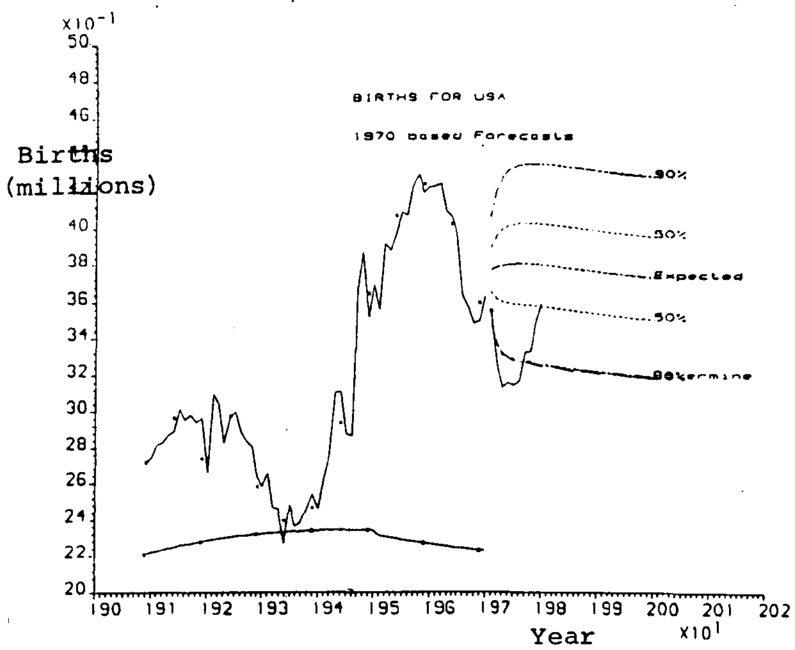




Forecasts are displayed for the USA in figure 4.3.7 for forecasts made in 1950, 1970 and 1980. 1930 is not used as a base for the USA since there is insufficient data in the birth series to allow ARIMA modelling. For forecasts made in 1950 observed births immediately "breakout" above the upper 90% confidence interval and remain out for more than fifteen years. Coupling this with the fact that the trend in births for the previous 10 years was upwards gives strong evidence of the need for escalating the model and that this was due to a sustained rise in fertility, since the other vital components did not change significantly over this period. When this escalation is incorporated into the model future births "fall out" of the bottom of the 90% confidence region on the 1970 forecasts indicating that one has "over corrected". When more information is available by 1980 the expected level of future births stays similar to that of the 1970 based forecasts, this is credible since the children born in the "baby boom" of the 1960's will be of parentable age and so by 1970 one should anticipate a rise in total births even if the overall trend in family size is downwards.

Figure 4.3.7 - U.S.A. Births





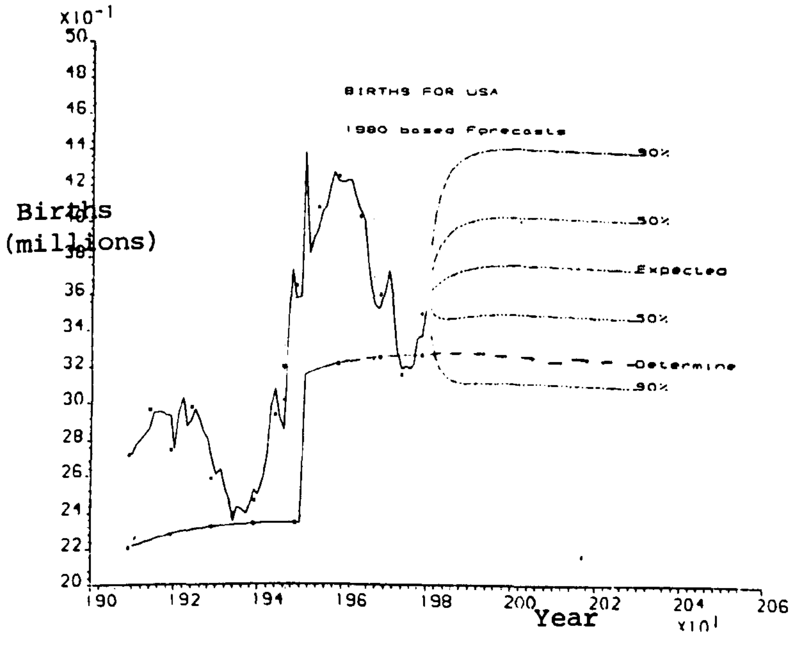


Figure 4.3.8 displays birth forecasts for France made in 1931, 1971 and 1980. The 1931 based forecasts display similar features to 1950 based forecasts for the USA in that by 1950 the observed births rise above the upper 90% confidence interval and generally stay above for a twenty year period, and since the deterministic trend is downwards one concludes that escalation is indeed required for France after 1950. Correcting for escalation improves forecasts and as a result by 1971 observed births are back in the confidence region. Forecasts made from 1980 indicate that births will fall but the confidence region is very wide (±40%) indicating the lack of precision.

Forecasts for Scotland based 1931 and 1981 are displayed in figure 4.3.9. The number of births by 1970 fall well below the lower 90% confidence interval and remain well below. Prior to 1970 the forecast interval has "captured" the future observed births which are scattered around the deterministic level which is at the lower end of the confidence region. This suggests that one has no reason to doubt the model up to 1970 however, descalation is required after this date.

Leach (1981) shows that de-escalation is required by 1940. This was chiefly due to high emigration and is not shown in the birth modelling until much later. Figure 4.3.9 shows that after making this correction the deterministic level of births falls, but the expected value of births suggests a rise and this is likely as it reflects the counter cyclical nature of fertility. (No 1971 based forecasts have been made, as not enough information was available to fit satisfactorily the escalated model).

Figure 4.3.8 - French Births

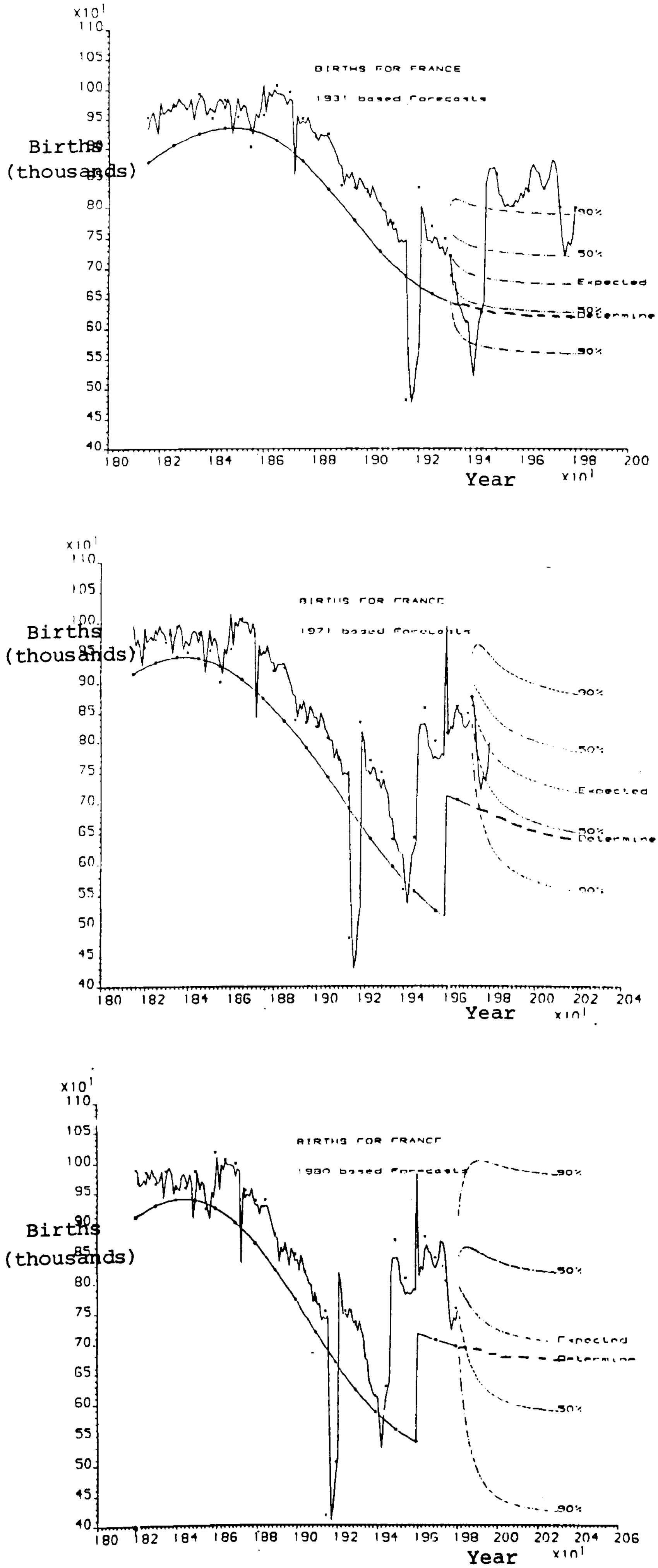
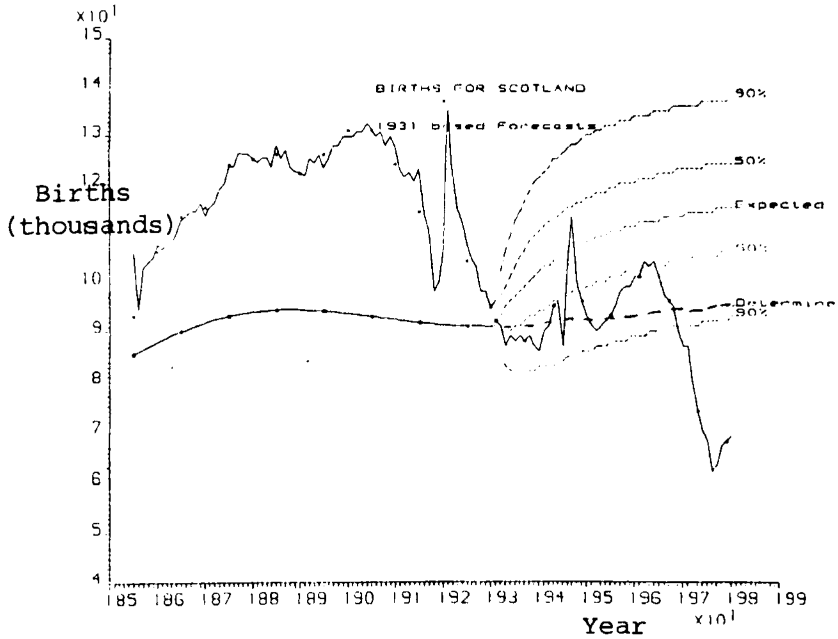
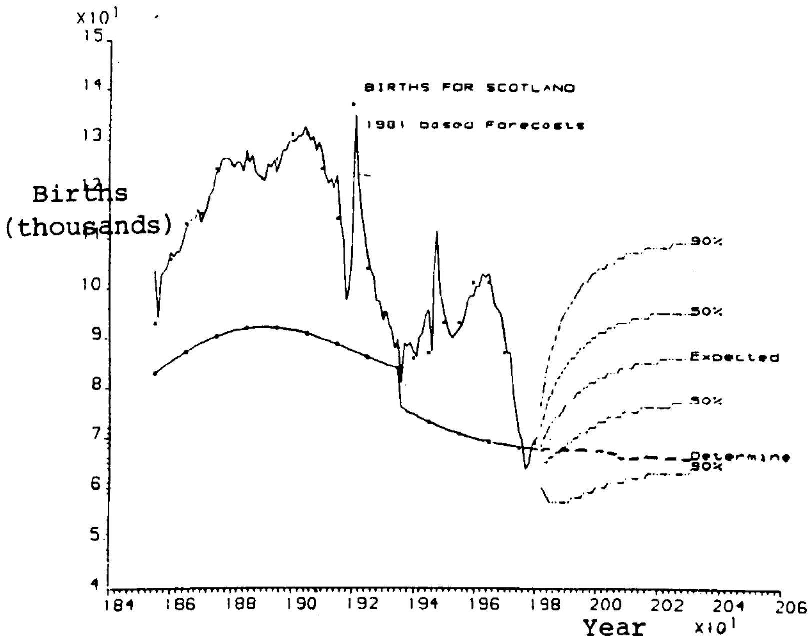


Figure 4.3.9 - Scottish Births



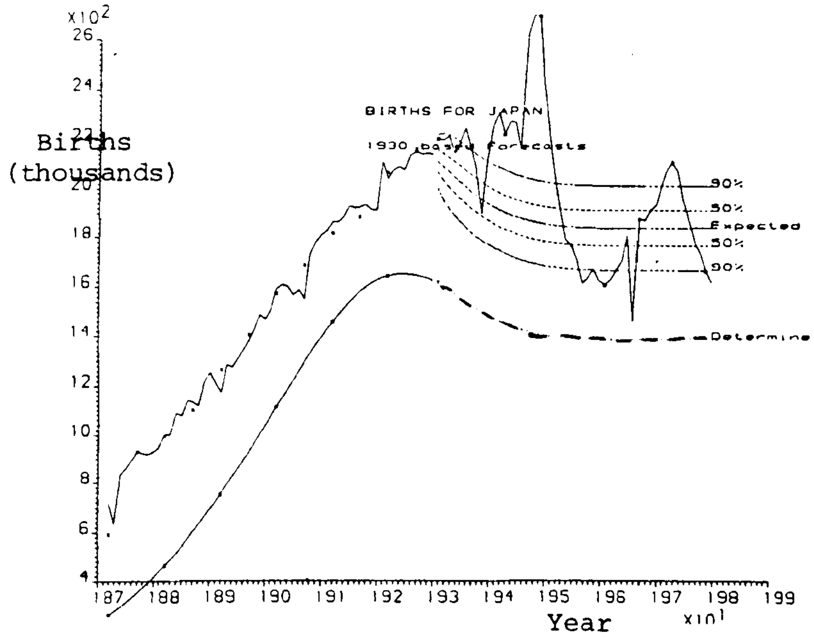


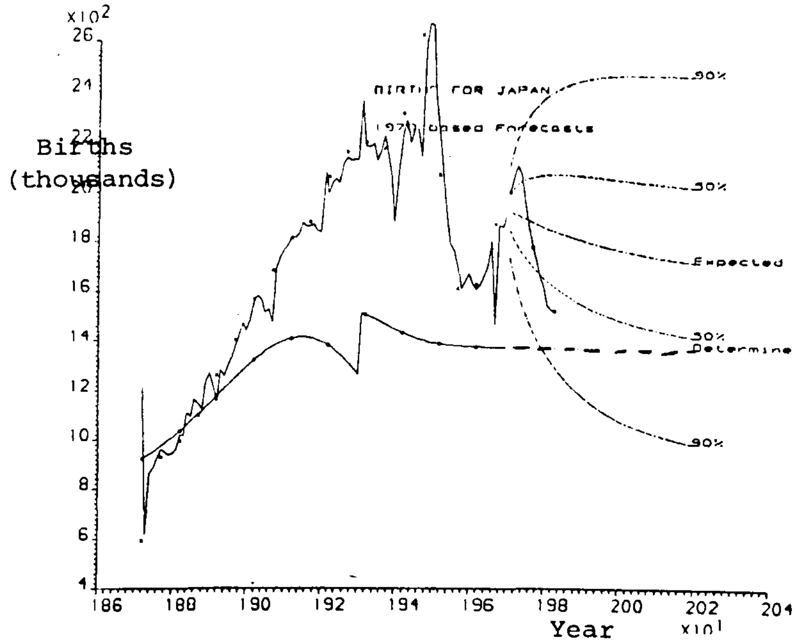
Forecasts of Japanese births are illustrated in figure 4.3.10 made from base years 1930, 1970 and 1983. For the 1930 based forecasts the deterministic level has been consistently low and future observed births soon break out of the upper end of the confidence regions which is compared to other regions, rather narrow. This suggests escalation which is incorporated into the 1970 based forecasts. These forecasts encompass all future births within the region. 1983 based forecasts show that the likely future of births is close to the deterministic level which is fairly flat and near replacement level.

The models which were used to generate the forecasts presented in figures 4.3.6 to 4.3.10 are laid out in table 4.3.5.

When deterministic models are made of births and subtracted from observed birth series a near stationary series is left, which can be modelled by low order ARMA models. These are mainly of an auto-regressive form indicating long term movement of the mean level towards the deterministic level and eventual replacement level fertility. This is consistent with using sigmoid to model total population for as the upper asymptote is approached if migration and mortality do not change drastically (as is likely to be the case for the developed nations) then births should approach an asymptotic level $\bar{\mu}$ years earlier.

Figure 4.3.10 - Japanese Births





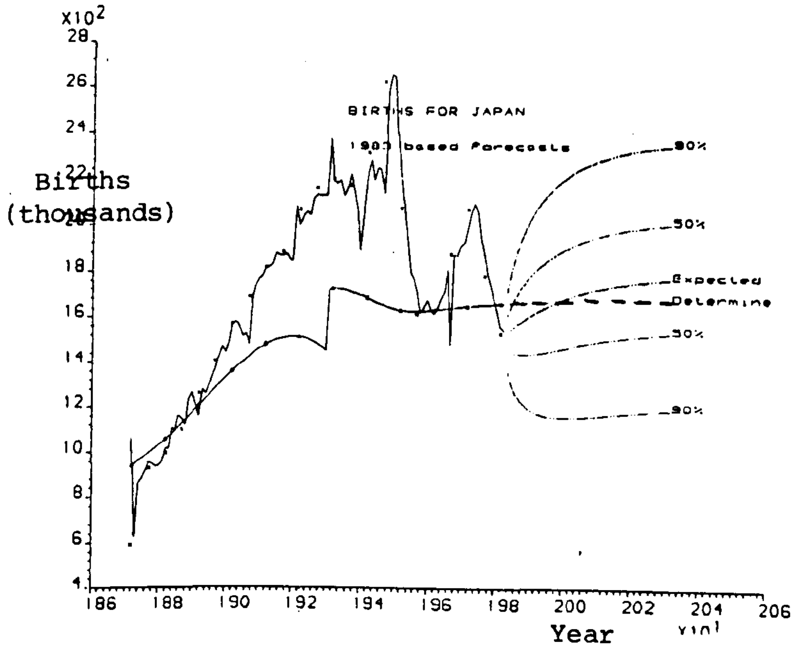


Table 4.3.5 Models of Birth Series (Births in thousands)

Birth Series	Deterministic Part: Population Logistic P_∞	P_0	c	d	Deterministic Part: e_{ot} Logistic E_∞	E_0	γ	ψ	Stochastic Residual Model
Scotland 1855-1931	6.80	1.608	0.017	—	24.70	0.055	0.0533	42.35	ARMA (1,0) $Z_t = 0.929Z_{t-1} + e_t + 20.66$ (0.053) (7.30)
Scotland 1855-1981	1951-1981 5.279	0.646	0.035	—	38.07	0.144	0.0423	41.90	ARMA (1,0) $Z_t = 0.936Z_{t-1} + e_t + 20.79$ (0.036) (6.20)
France 1815-1931	41.73	27.23	0.0213	—	29.11	0.08	0.0538	38.15	ARMA (1,0) $Z_t = 0.729Z_{t-1} + e_t + 54.49$ (0.064) (15.95)
France 1815-1971	1950-1971 65.50	42.74	0.0012	—	55.06	0.653	0.032	35.68	ARMA (1,0) $Z_t = 0.865Z_{t-1} + e_t + 79.02$ (0.04) (28.27)
France 1815-1980	1950-1982 57.28	0.0019	0.732	—	48.75	0.431	0.036	36.44	ARMA (1,0) $Z_t = 0.908Z_{t-1} + e_t$ (0.043)
Sweden 1875-1930	9.73	1.17	0.014	—	46.17	1.36	0.0222	33.89	ARMA (1,0) $Z_t = 0.913Z_{t-1} + e_t - 27.76$ (0.064) (6.45)
Sweden 1875-1970	13.24	1.33	0.0116	—	66.42	1.59	0.0198	32.01	ARMA (1,0) $Z_t = 0.971Z_{t-1} + e_t$ (0.029)

Table 4.35 - Continued

Birth Series	Deterministic Part: Population Logistic P_∞	P_0	c	d	e_{ot} Logistic E_∞	E_0	γ	Ψ	Stochastic Residual Model
Sweden 1875-1980	13.21	1.31	0.118	—	68.00	1.786	0.0191	33.40	ARMA (1,0) $Z_t = \underset{(0.031)}{0.987}Z_{t-1} + e_t$
U.S.A 1909-1950 births in millions	193.79	3.898	0.0315	—	46.58	4.66	0.201	30.29	ARMA (4,0) $Z_t = \underset{(0.158)}{1.326}Z_{t-1} - \underset{(0.247)}{0.763}Z_{t-2} + \underset{(0.251)}{0.694}Z_{t-3} - \underset{(0.184)}{0.346}Z_{t-4} + e_t + \underset{(0.23)}{0.60}$
USA 1909-1970 births in millions	1950-1970 262.70	150.697	0.0019	—	123.98	17.05	0.0117	5.49	ARMA (1,0) $Z_t = \underset{(0.051)}{0.922}Z_{t-1} + e_t + \underset{(0.02)}{0.96}$
USA 1909-1980 births in millions	1950-1980 270.78	0.189	0.047	—	68.47	2.23	0.0295	15.77	ARMA (1,1) $Z_t = \underset{(0.065)}{0.862}Z_{t-1} + e_t + \underset{(0.118)}{0.396}e_{t-1} + \underset{(0.16)}{0.52}$
Japan 1872-1930	32.04	0.0074	0.076	33.77	54.68	17.13	0.002	25.99	ARMA (1,0) $Z_t = \underset{(0.086}{0.787}Z_{t-1} + e_t + \underset{(37.05)}{451.66}$
Japan 1872-1970	90.52	22.58	0.0072	33.94	37.06	0.0001	0.120	43.80	ARMA (1,0) $Z_t = \underset{(0.032)}{0.971}Z_{t-1} + e_t + \underset{(299.48)}{288.33}$
Japan 1872-1983	1922-1983 70.89	0.0005	0.073	56.38	36.76	0.000001	0.112	43.62	ARMA (1,0) $Z_t = \underset{(0.033)}{0.954}Z_{t-1} + e_t + \underset{(187.15)}{119.79}$

Regarding the quality and validity of these forecasts one should be sceptical about the practical worth and likelihood of the upper asymptote when related to something as fickle as numbers of human births per year. Conclusions drawn so far are:-

i. The confidence region is very wide - thus precision about the future of births is low.

ii. The deterministic level of births is on occasion far from the true level and when departure becomes large and consistently positive or negative at times close and prior to the base then there are indications of a change in the phase of growth. This is confirmed if future births breakout of the confidence region for a sustained period.

iii. The short term forecasts (up to five years) for cases where there have been no changes in the pace of growth, have been good and in the correct direction, but the ARMA models are of too low order to reflect the cyclical nature of the birth series, in long term forecasts.

iv. The expected path is always in the correct direction if one considers the birth series to be a "counter-cyclical" process. That is if a trough had been observed and the most recent births were higher than this then the expected path would show a downturn.

From this one can form a decision rule which will aid in the early detection of a change in the phase of growth and allow corrective measures to be made. The rule is:-

If there are no significant changes in migration or mortality and the following are true:-

1. The deterministic path over the last, say ten years, has been at least ten percent consistently high or low.

2. The future births "break out" of the 90% confidence region for a sustained period (arbitary set at 10 years).

3. The expected ARMA path fails to move in the direction of the trend of births over the next ten year period.

Then escalation or de-escalation is expected. If only 1 and 2 hold true a change is probable and if only 1 holds exercise caution.

Having detected the need for escalation then one should switch to a simpler way of modelling population. Murphy's (1980) asymptotic method for modelling total population is suggested which can be used to generate births by a residual procedure in conjunction with mortality (D) and migration (M)

$$\text{ie} \quad B_{t-1} = (P_t - P_{t-1}) + D_{t-1} - M_{t-1}$$

This will allow forecasts to be made without having to rely on the past outdated data which belonged to a different phase and should be used until sufficient information becomes available for more complex modelling.

To assess further the quality of the forecasts the mean absolute percentage error of the forecasts made for ten years ahead which are based in 1930/1 (1950 in the USA case) and 1970/1 are displayed in table 4.3.5

Table 4.3.5 - Ten year MAPE for birth forecast

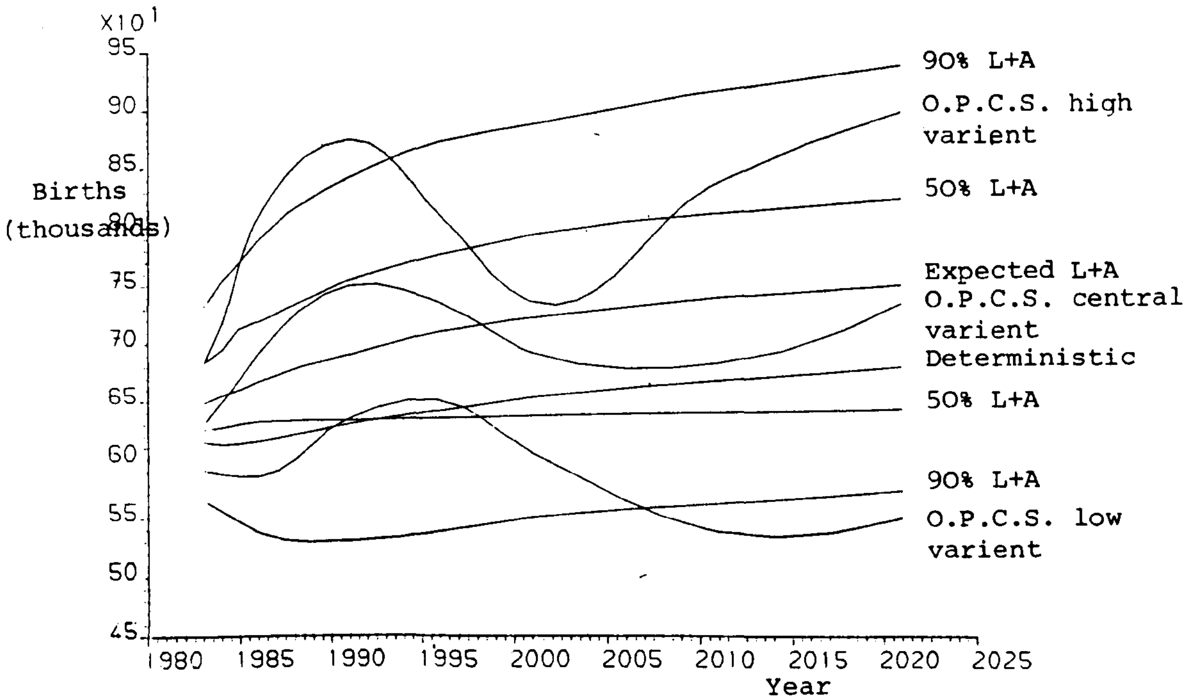
Nation	Base	MAPE
England and Wales	1931	8.86
	1971	23.77
Scotland	1931	2.97
USA	1950	30.09
	1970	14.98
France	1931	12.51
	1971	5.71
Sweden	1930	3.70
	1970	5.95
Japan	1930	6.01
	1970	7.45

Excepting cases where escalation is detected (ie France and USA) and the English and Welsh 1971 case (which was soon put "right" with updated information) forecasts of births are satisfactory.

Thus the method of using an ARMA model with a logistic trend derived from the model of total population can generate useful forecasts of births (which, at least for England and Wales, compare favourably with official forecasts). Forecasts of births should be carried out if total population is to be forecasted as it allows the early detection of a change in the phase of growth and allows appropriate remedial action to be identified.

Considering again, the wide confidence limits that are given, but as can be seen from figure 4.3.11, which displays 1981 based forecasts and 1981 based OPCS projections (OPCS, 1984), although not reflecting cyclical movements in births the Logistic - ARMA forecasts encompass virtually the entire span of the OPCS variant projections. Therefore covering the possible future spread under practical assumptions thus the statistical confidence limits are probably of more use to the user since - "there is no way of assigning measures of probability to the variant projections". (CPRS 1977), and the width of the limit highlights the imprecision inherent in forecasting births. Part of this imprecision comes from the mathematical limitations of the models to model accurately past fluctuations and part as a result of the manner by which factors conspire to influence future births likely being novel. Since the probability of a family bearing X number of children is determined by a complex two way relationship with a continuously evolving system that

Figure 4.3.11 - O.P.C.S. Projections & Logistic Based Forecasts



L+A = Logistic + ARMA forecast

involves economic physical, social and psychological factors. The decision making process and outcome is difficult to predict. This is discussed further in chapters 5 and 6.

Also as Hajnal (1955) points out the production of a forecast may in itself alter the course of births for example if below replacement fertility is predicted for a country then the Government of that country may set in motion a vigorous pro-natal policy leading to an upsurge in the number of births (this may account in part for the French 1945-50 escalation, which is discussed further in Chapter 6).

Another useful feature of this logistic-ARMA modelling is that forecasts are easily updated with the latest birth count which updates the ARMA part and the deterministic part is updated when the next census is enumerated.

Finally one other observation of this birth modelling is that when the auto correlation structure of the birth series minus the deterministic model is examined a significant negative peak autocorrelation is observed between lags 27 - 32. This reflects the generation effect on births. However, when this was incorporated into the ARMA models by introducing a 'seasonal' effect no significant difference was observed and parameter estimation was harder. Therefore the simpler modelling is preferred.

Unfortunately the use of this model is limited for nations which have limited birth series data which precludes the use of ARMA models. However, the deterministic trend can still be extrapolated which for many nations where there are limited data may well provide satisfactory forecasts. This is particularly useful for nations which have not yet undergone recent fertility transition, since the greatest departure from the deterministic path is associated with post-transition births. In all cases the deterministic trend is likely to be the best long term forecast and the expected ARMA path, where obtained, is likely to be the best short term forecast.

4.4 Modelling and Forecasting the Number of Deaths per Year

Time series of the total number of deaths per year are now examined. Non-age specific measures of the number of deaths per year are usually of little interest to "bottom up" demographic modellers, however in this analysis forecasts of these are made.

Applying ARIMA modelling to the series of deaths per year for England and Wales over the period 1838 to 1981 (144 observations) gives an AR(2) model, but the spread of residuals are wide, there are non stationary problems and the correlogram indicates that transformation of the series is required. The most suitable transformation appears to be to take natural logarithms of the series.

The models arrived at are:-

untransformed series

$$Z_t = \underset{(0.08)}{0.58}Z_{t-1} + \underset{(0.08)}{0.39}Z_{t-2} + e_t + 505.54$$

ln transformed series

$$\ln Z_t = \underset{(0.08)}{0.60} \ln Z_{t-1} + \underset{(0.08)}{0.39} \ln Z_{t-2} + e_t + 6.36$$

Unfortunately stationarity has still not been obtained. Better results were obtained when the crude death rate series is used. The crude death rate series for England and Wales was first modelled over the period 1838 to 1901 and repeated by extending by ten years up to 1981. From these models forecasts were made upto 1981 and compared to observations and 95% confidence limits are generated. Table 4.4.1 displays these models. It was observed that for these models stationarity was achieved and that the forecasts made from these models are good even for fifty year ahead forecasts. (The mean percentage error in 1931 based forecasts over the period 1932 to 1981 is only around 5% most of the predictions erring on the high side. 1970 based forecasts covering the period 1972 to 1981 have a MAPE of only 1.63%).

From table 4.4.1 one can observe that the models are stable over time all models being of the same order. The parameter of the moving average term is fairly consistent ranging from 0.520 to 0.648 having a mean of 0.570.

Table 4.4.1 - ARIMA Models - For English and Welsh Death Series

Period	Model	Sum of Squares	Eventual Forecast Value
1838-1901	$Z_t = Z_{t-1} + e_t - \underset{(0.097)}{0.648} e_{t-1}$	71.815	17.792
1838-1911	$Z_t = Z_{t-1} + e_t - \underset{(0.104)}{0.520} e_{t-1}$	80.759	14.333
1838-1921	$Z_t = Z_{t-1} + e_t - \underset{(0.096)}{0.548} e_{t-1}$	98.517	12.597
1838-1931	$Z_t = Z_{t-1} + e_t - \underset{(0.087)}{0.554} e_{t-1}$	104.607	11.986
1838-1941	$Z_t = Z_{t-1} + e_t - \underset{(0.082)}{0.564} e_{t-1}$	108.659	12.617
1838-1951	$Z_t = Z_{t-1} + e_t - \underset{(0.078)}{0.571} e_{t-1}$	113.194	11.893
1838-1961	$Z_t = Z_{t-1} + e_t - \underset{(0.074)}{0.574} e_{t-1}$	114.181	11.711
1838-1971	$Z_t = Z_{t-1} + e_t - \underset{(0.071)}{0.579} e_{t-1}$	118.175	12.261
1838-1981	$Z_t = Z_{t-1} + e_t - \underset{(0.069)}{0.573} e_{t-1}$	115.768	11.849

There are some instabilities in the series however, such as around the war years. Perhaps this period should be treated for outliers. If this had been done then the width of the confidence limits, which are wide (± 25%), might be reduced.

Similar models were also found for the USA, Sweden and France. Some of these models are displayed in table 4.4.2, along with their eventual forecast function.

These models fitted well and produced satisfactory forecasts. The MAPE for the most recent ten year forecasts are as laid out in table 4.4.3.

In all cases the best model is when the series is differenced once and modelled as a moving average process. The optimal model for all cases appearing to be an ARIMA (0,1,1) model which is clearly defined when the autocorrelations and partial autocorrelations are examined. This model has an eventual forecasting function of:-

$$F(t) = F(t-1)$$

ie. the next periods forecasts is the same as the previous forecasts so the forecast level is constant. This suits well the situation in developed nations where currently there is little trend in year to year variation in crude death rates and is sensibly defined as a random walk.

Table 4.4.2 - Models of Crude Death Rates

Series	Model	Residual sum of squares	Eventual Forecast Function
USA * 1920-1970	$Z_t = Z_{t-1} + e_t - \underset{(0.132)}{0.395}\, e_{t-1}$	4.451	9.51
USA * 1920-1979	$Z_t = Z_{t-1} + e_t - \underset{(0.124)}{0.354}\, e_{t-1}$	4.750	8.79
Sweden 1875-1970	$Z_{t=Zt-1} + e_t - \underset{(0.095)}{0.379}\, e_{t-1}$	175.524	14.66
Sweden 1875-1980	$Z_t = Z_{t+1} + e_t - \underset{(0.091)}{0.379}\, e_{t-1}$	176.139	15.49
France 1801-1969	$Z_t = Z_{t-1} + e_t - \underset{(0.059)}{0.652}\, e_{t-1}$	499.803	11.16
France 1801-1979	$Z_t = Z_{t-1} + e_t - \underset{(0.057)}{0.651}\, e_{t-1}$	500.679	10.28

*This series started in 1909 but modelling only began since 1920 in order to avoid instabilities around the first world war and a particularly high peak in 1918 associated with an influenza epidemic.

Table 4.4.3 - M.A.P.E. for short lead time forecasts of death rates - Forecasts based circa 1970

Population	5 year lead time	10 year lead time
England and Wales	2.22 L	1.63 L
USA	3.01% H	6.34% H
Sweden	2.98% L	4.07% L
France	5.05% H	6.60% H

L = Forecasts are predominantly low
H = Forecasts are predominantly high

Many developing nations are exhibiting a clear trend in declining death rates as life expectancy at birth increases rapidly (see Gwatkin and Brandel, 1982) and so more complex models are likely to be required. For these countries an approximate estimate of the magnitude of the number of deaths per year (D_t) can be obtained from the change in population size minus the number of births, so ignoring the effects of migration ie:-

$$D_t = P_{t-1} - P_t + B_t$$

In this case for developing nations P_t it is suggested be modelled by Murphy's (1980) recursive asymptotic method described in Chapter 3 Section 3.5 and B_t by the method discussed in Section 4.3 ie:-

$$B_t = \frac{P(t + \frac{1}{2}e_o)}{e_o}$$

4.5 Modelling and Forecasting Migration

In this section models of migration are considered only allowing knowledge on past migration levels, other vital components and the total population size as data inputs.

To model Swedish immigration, emigration and net migration over the period 1875 to 1960 and make forecasts, it was found best to use crude rates rather than numbers. The models fitted were:-

(a) Immigration

ARIMA (3,1,0) $$\underset{}{(1-\underset{(0.11)}{0.09B} + \underset{(0.11)}{0.15B^2} + \underset{(0.11)}{0.18B^3})(1-B)Z_{t-1}=e_t}$$

model specification is poor.

(b) Emigration

ARIMA (2,0,0) $$Z = \underset{(0.106)}{1.027Z_{t-1}} - \underset{(0.106)}{0.182Z_{t-2}} + e_t$$

(c) Net Migration

ARIMA (2,1,0) $$(1-\underset{(0.105)}{0.111B} + \underset{(0.105)}{0.253B^2})(1-B)Z_t = e_t$$

For all models specification is poor.

MAPE's for the twelve year forecasts are:-

Immigration = 27.10%

Emigration = 17.61%

Migration = 109.73%

observed migrants versus forecasted immigration - forecasted emigration = 79.62%

Thus ARIMA time series modelling is not satisfactory and shall be pursued no further.

A more suitable approach to obtain the scale of migration might be to use a residual method, suitably adapting the balance equation ie:-

$$M_t = (P_t - P_{t-1}) - B_t + D_t$$

The level of net migration has been calculated in this way and compared to observed net migration over a ten year period 1971-1980 for the population of Sweden and U.S.A.

Surprisingly the results when compared to observed net migration proved to be unsatisfactory, the MAPE's for the period for both populations was in excess of 60%. This result arose despite adequate population, birth and death forecasts. The conclusion then is that alternative modelling for migration is required, perhaps causal econometric modelling would be more suited.

4.6 Age Specific Models of Births

The aim of age specific birth modelling is to model births by age of mother and use this model to make predictions. This is only of use if changes in the age structure of the population at risk of giving births is reflected in the modelling. Unfortunately this is not possible by following a top down approach and simply disaggregating total births. Age specific forecasts must first be obtained. (This is discussed in the next section). However, for near stationary populations disaggregating total births on a proportionate basis might give useful information. In this section such an approach applied to English and Welsh age specific births is presented.

The data used is the number of births recorded by five year age groups of mother at the year t (b_{at}, where a represents the age groups :- 15-19, 20-24, - - - 45-49. This data is then expressed as a proportion of total births at time t(b'_{at})).

Attempting modelling by taking each age group separately, by for example ARIMA techniques is unsatisfactory as it fails to account for the inter-relationship between each b'_{at} series. For example if births in the age group 20-24 are high at time t-5 then one would expect that births to the age group 30-34 at time t+5 would be low as this cohort of mothers complete their child bearing earlier in life.

Methods of time series analysis that can be used for interdependent series are being developed, such as MARIMA (see Cox et al, 1983). But such methods have not been successful for more than three series and require substantial data and computing power and so this type of modelling is as yet unsuitable for use in demography.

Since b'_{at} can be thought of as a probability distribution then an appropriate modelling procedure would be to fit a probability distribution function to b'_{at} for the different years, and model how this function changed.

A preliminary investigation of fitting probability distributions to b'_{at} over the period 1938 to 1981 for England and Wales, using the E.C.S.L. computer package consistently gave the three parameter Weibull distribution as the closest approximation to b'_{at}. (The E.C.S.L package automatically fits most of the standard statistical distributions, (E.C.S.L./C.A.P.S, 1980)

The Weibull distribution is described as:-

Variate W: l,s,c Range $0 < X < +\infty$

l = location parameter

s = scale paramter (b>0)

c = shape parameter (c>0)

Cumulative Distribution function $F(x) = 1-e^{-((x-l)/s)c}$

(Although unbounded at the upper end, the error of allowing an infinite tail is slight).

It is not surprising that one should find the Weibull to be the most suitable distribution as it is extremely flexible. To verify the suitability of the Weibull and to obtain better estimates of the parameters the distribution is fitted to the observed data using a FORTRAN IV programme based on the non linear least squares NAg10 Library routine E04FDF.

The fitting algorithm is:-

Minimise $$Z(x) = \sum_{i=1}^{m} [F(x_{i-1}) - F(x_i)]^2$$

where m = number of age groups = 7

x_i = 19.5, 24.5, - - - -, 49.5

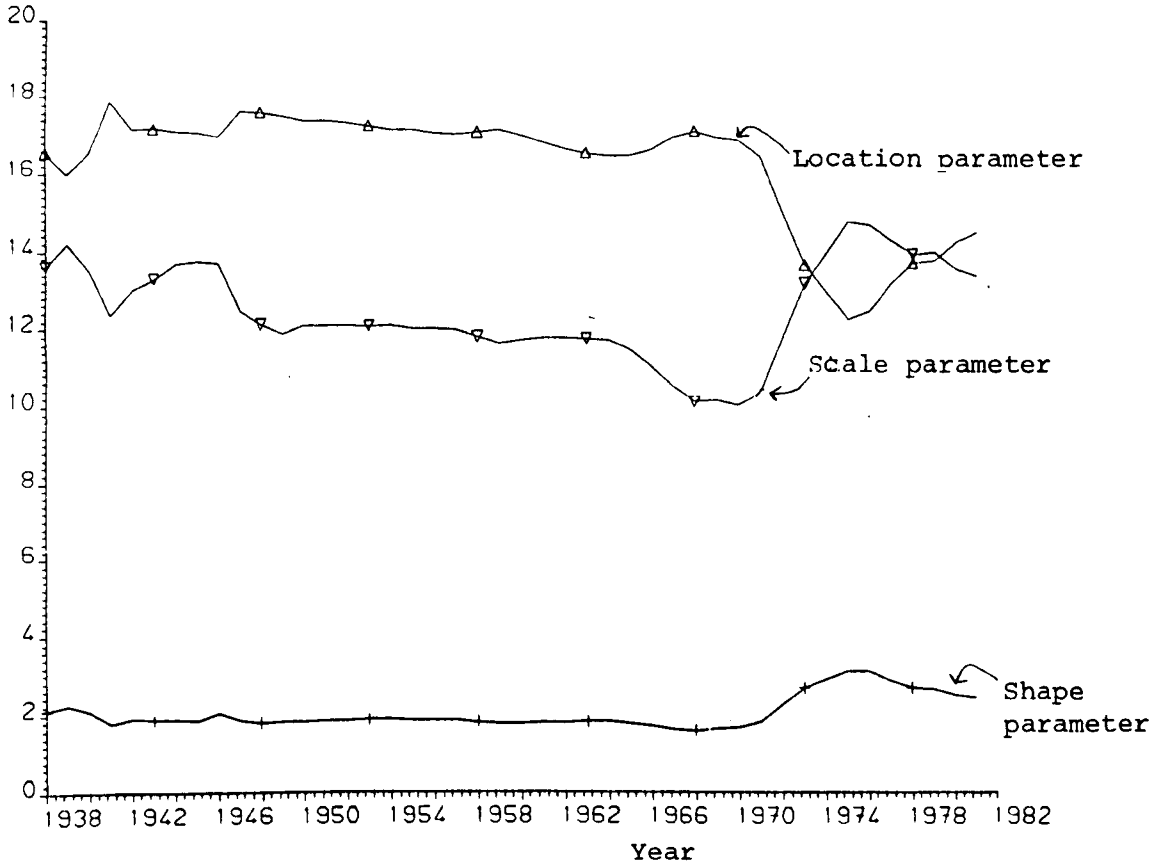
The results from this analysis confirm the Weibull to be an acceptable model. The variation of the parameters with time are displayed in figure 4.6.1.

Up to 1970 the parameters appear stable, but after 1970 there is a sudden departure from trend as the distribution changes from a location just above 16 years to less than 14 years. This is unfortunate and casts doubt on the forecasting ability of such a procedure.

It is sensible to fix the location of the Weibull to prevent it falling below 16 years. This parameter is fixed at 16 years, the mean of the observed level of this parameter. The Weibull is refitted and the shape and scale parameters noted.

On doing this the shape parameter appeared constant, having a mean of 2.001 and variance of 0.0142. Therefore it might be reasonable to fix this parameter at 2.0.

Figure 4.6.1 - Variation in the Parameters of the Weibull



The Weibull is refitted with the two parameters held constant. The scale parameter (S_t) is modelled over time - a suitable model is a modified exponential parameterised as:-

$$S_t = \underset{(0.97)}{9.72} + \underset{(0.83)}{4.96}e^{\underset{(0.013)}{-0.033t^*}} \qquad \text{where } t^* = \text{year} - 1938.$$

Each time a parameter was held constant the sum of squares of the residuals increased. When the location parameter was fixed the root mean square error (R.M.S.E) of all the fits rose from 8.452×10^{-3} to 11.339×10^{-3}. When the shape parameter was fixed the R.M.S.E. rose to 14.142×10^{-3}. This appears to be a linear rise and not to be an excessive cost for the benefit of the resulting simplification.

The fit of this "constrained" Weibull to the 1981 birth distribution is illustrated in figure 4.6.2. The model is used to extrapolate the distribution of γ_{at} into the future. (The 1981 based forecast of the birth distribution for 1986 & 1991 is illustrated in figure 4.6.3.) These forecasts are given high and low variants by inputing the 90% confidence limits on B_t.

The general trends suggested in these forecasts are that in England and Wales the proportion of births will be increasingly concentrated into the younger age groups and there will be an overall fall in fertility. This type of prediction is in agreement with the New Home Economics (NHE) model and of Ermisch's (1983) version of the NHE models for England and Wales.

Figure 4.6.2 - Fit of the Weibull Distributions

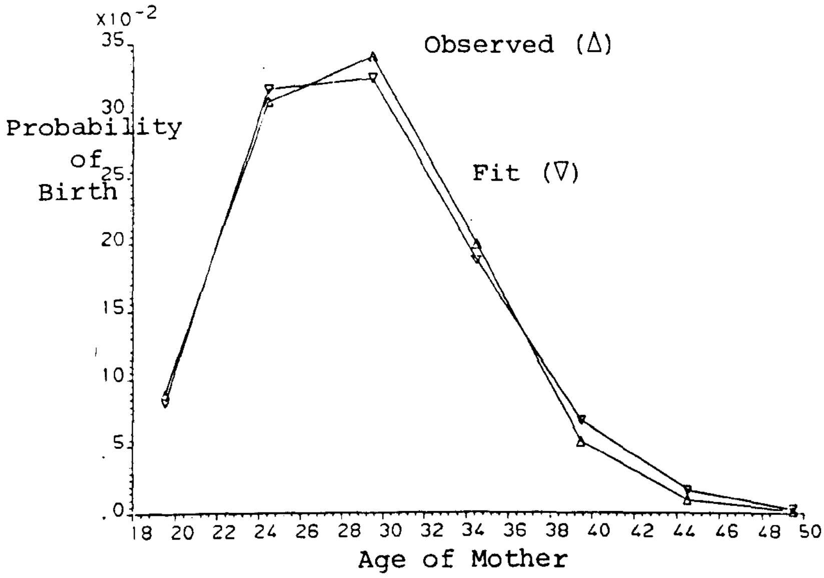
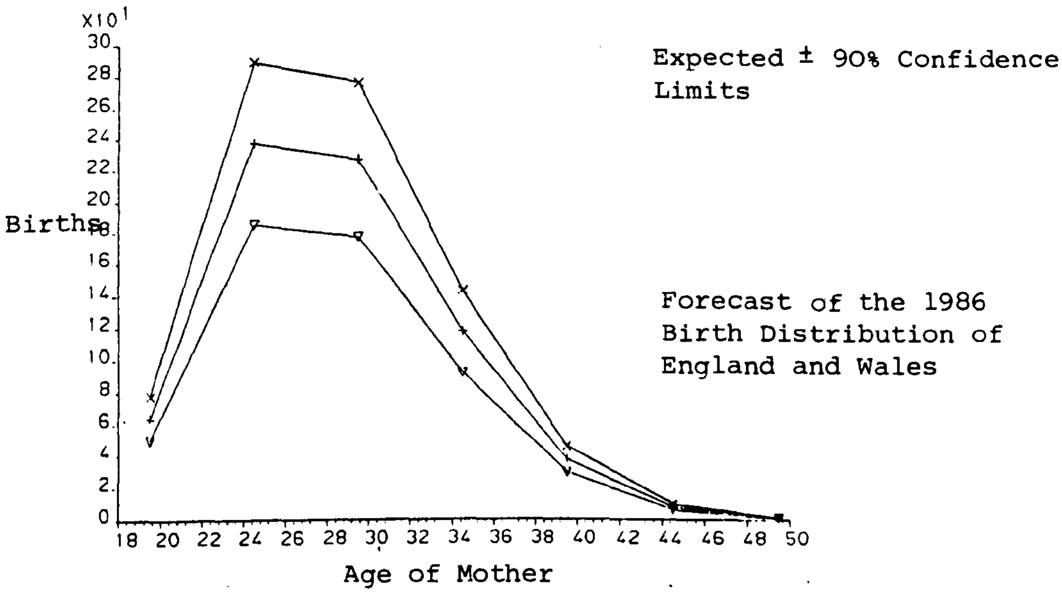
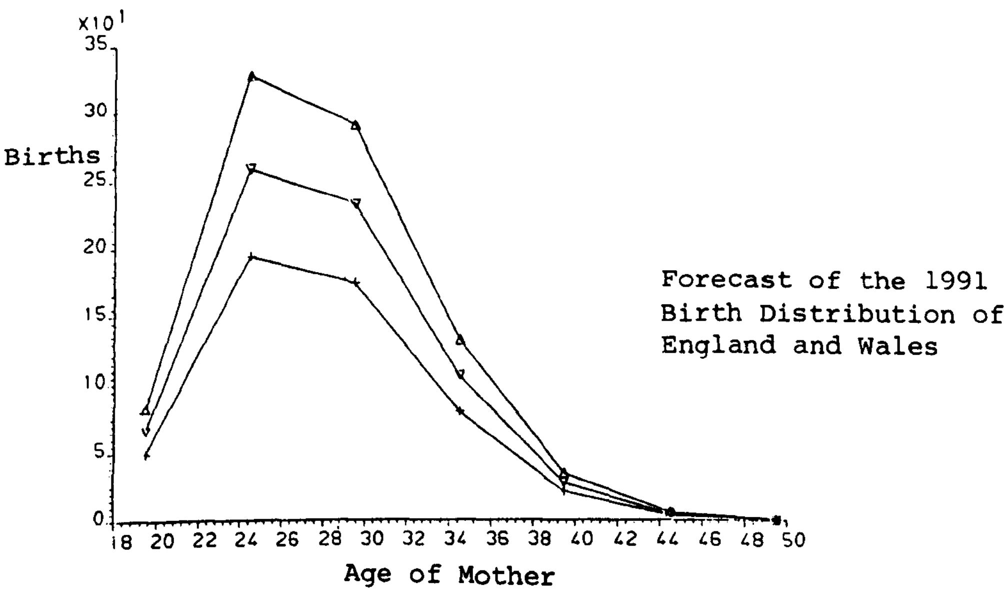


Figure 4.6.3 - Forecasts of "Weibull" Distributed Births (1981 based)





Forecasts of b'_{at} made for models fitted over the period 1938 to 1971 are compared to observations for the years 1976 and 1981. The model for the scale parameter for these forecasts is

$$b'_{\alpha t} = -14.53 + 29.18e^{-0.0042t^*};$$

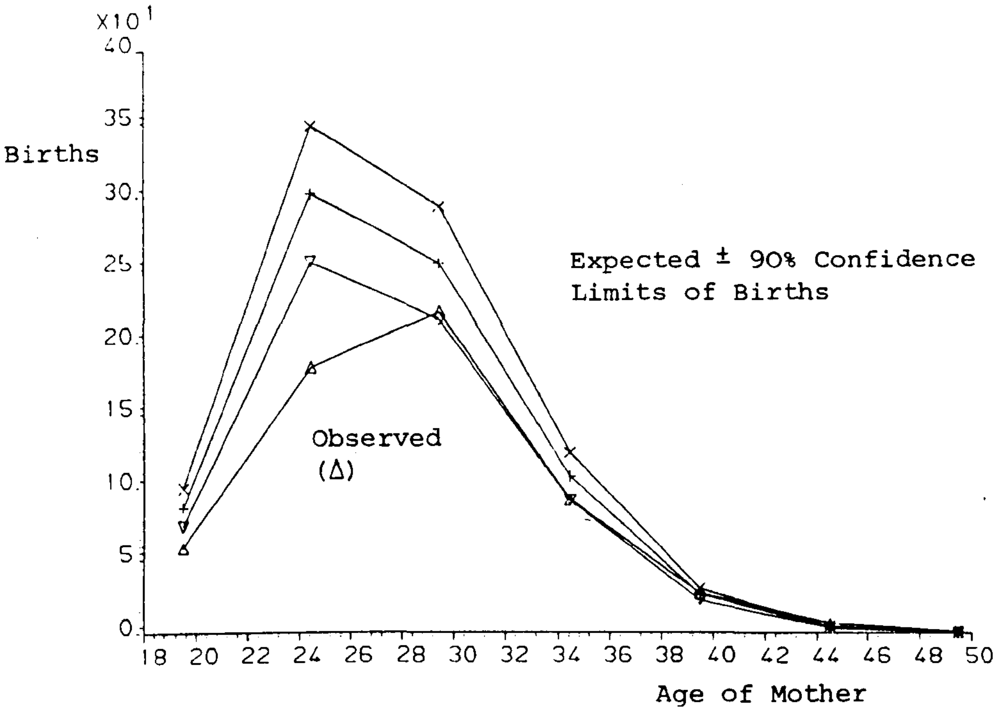
which is an unsuitable model on theoretical grounds as it is asymptotically bound at a negative value.

The forecasts for 1976 and 1981 are displayed in figure 4.6.4 and are disappointing reflecting two inadequacies.

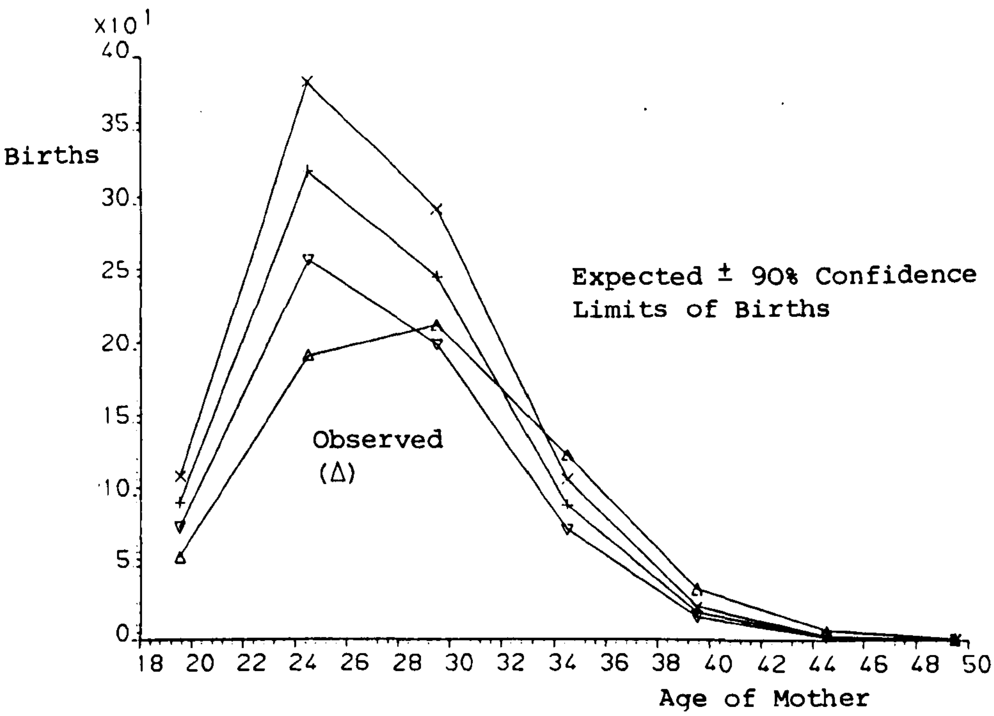
The forecasts are dependent on good forecasts of total births (these were poor for this period and the model is too simplistic to be able to incorporate deliberately delayed fertility. The Weibull model is too simplistic as it has continued the trend of the pre 1970's of higher fertility at younger ages failing to detect the movement to delayed fertility after 1970s. Thus to develop a modelling/forecasting scheme along the above lines more data is required in order to detect cyclical changes in fertility resulting from changes in cohort since marriage rates, female work participation, income changes and the like. (Existing models of fertility such as those developed by Butz and Wand 1979 or Ermisch 1982 are preferred to an elaboration of the above).

Figure 4.6.4 - Forecasts 'Weibull'
Birth Distribution (1971 based)

1976 Forecast



1981 Forecast



The method based on the Weibull distribution of forecasting b'_{at} might be improved if the parameters are not extrapolated by mathematical models but by using judgement made after considering changes in marriage patterns and economic factors.

4.7 Forecasting Age Structure

To model and forecast the sex-age structure of a population, trend models were fitted to the numbers in each five year age group over the period 1841 to 1981 for the English and Welsh population. Generally the four parameter logistic models were found to be the best descriptors of these series.

Unfortunately these models fail to reflect fluctuations in age groups which result from fluctuations in fertility and a vast amount of available information, namely those alive in the previous age group is not used. Hence, such an approach must be regarded as conceptually inadequate. It was also found that forecasts of age structure made by the trend models are inferior to simply assuming continued current rates of increase. (Similar inadequaces of such an approach were found by Murphy, 1980).

Since total births at year t (B_t) can be forecast then if one can project the conditional probability of surviving to age "a" given that the person was alive at t-1, at time t ($P_t(a)$), then one can construct the future age structure, assuming migration has little effect.

ie. $N_{ot} = B_t$

$N_{at} = N_{a-5,t-5} \times P_t(a)$ (taking 5 year age groups)

Where N_{at} is the size of the population in age group a at time t. Hence to forecast N_{at} one has first to forecast $P_t(a)$.

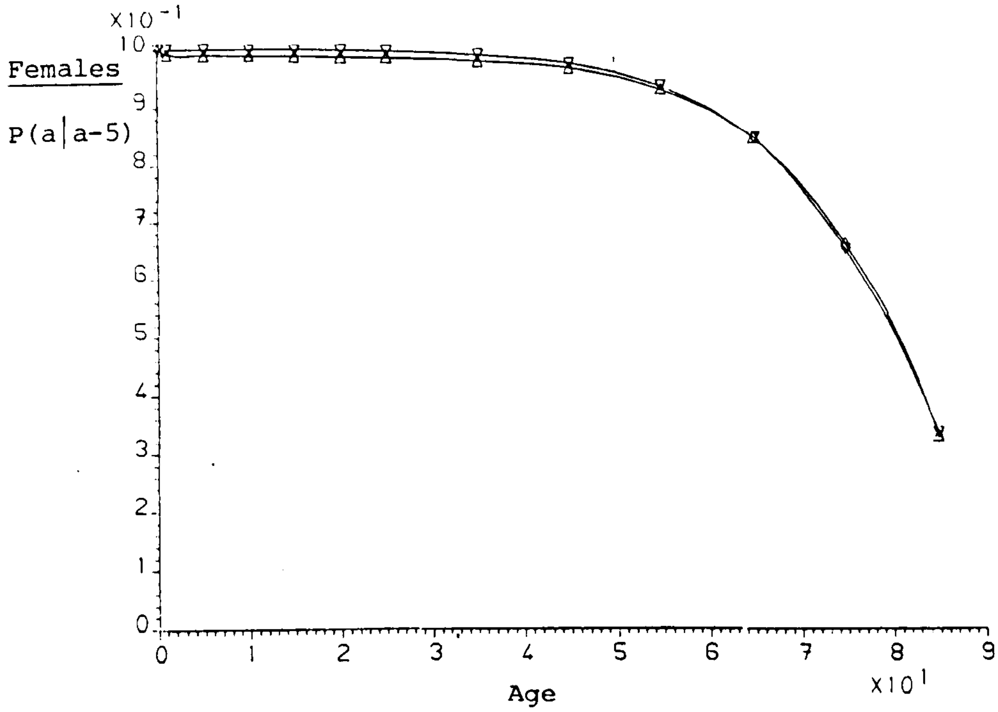
A series of observed $P_t(a)$ between 1951 and 1982 was computed for English and Welsh females and males. The Reed-Merrill transformation (see Section 4.3) was used to convert age specific mortality rates (${}_nM_a$) into P(a)(The conditional probability of surviving to a given the person was alive at a-n).

Survival functions (see Elandt - Johnson and Johnson, 1980) express the probability of surviving to age "a" as a function S_a of age. Of this class of functions the Gompertz survival function:-

$$S_a = P_t(a) = e^{R(1-e^{ca})}$$

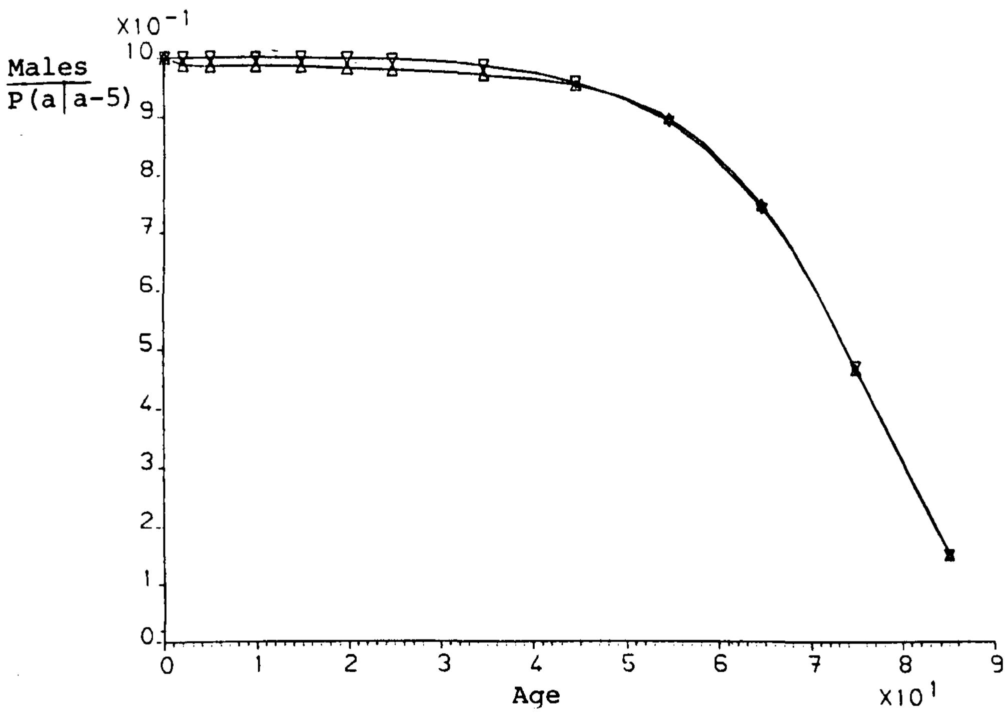
gives a suitable model of the English and Welsh data when fitted by least squares methods to $P_t(a)$. Theoretically this should be fitted to the survival probabilities of a cohort. However, since the population of England and Wales is approaching stationarity then S_{at} should give a reasonable approximation of $P_t(a)$. The fit of this function to the 1981 values of P(a) for males and females is illustrated in figure 4.7.1. The variation in the sum of squares of the residuals when S_{at} is fitted

Figure 4.7.1 - Survival Functions Fitted to Observed Probabilities for England and Wales - 1981



∇ = fit

Δ = observed



over the period 1951 to 1982 are displayed in figure 4.7.2 from which one can see that the fits are exhibiting the desirable property of getting better with time. (This is mainly as a result of decreasing infant mortality).

The Gompertz survival function does not adequately model the initially higher mortality associated with infancy and it is best only to fit the age groups greater than five years old and assume linear interpolation between S_0, which is equal to one, and the value of S_5.

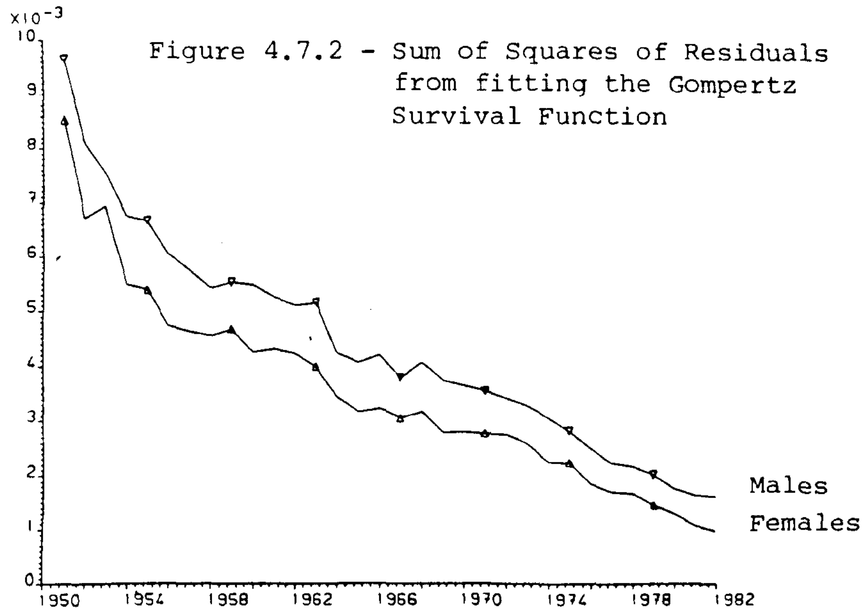
To extrapolate S_a into the future the parameters R and C are made time dependent. Models of R and C are formed from fitting the values of R and C recorded over the period 1951 to 1982. The variation of R and C over this period is displayed in figure 4.7.3.

The models of these parameters are laid out in table 4.7.1. (Also displayed are models fitted over the period 1951 to 1971, which will allow forecasts to be compared with observed values for the period 1972 to 1982).

The models outlined in table 4.7.1, apart from C_{tf} (1951-1971) are all bounded and suggest that with increasing t that S_a will approach a "sensible" non-zero asymptote.

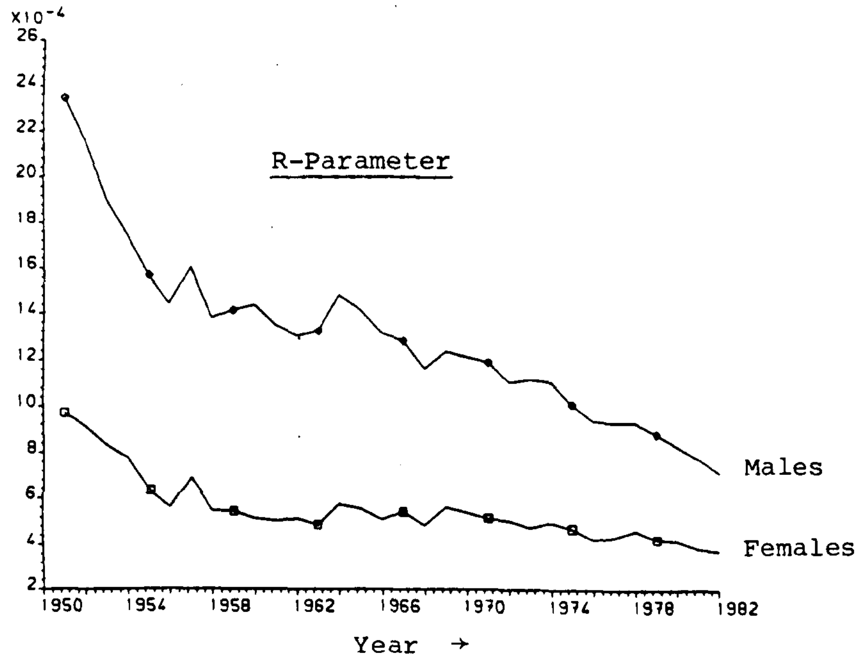
For practical purposes, since, as can be observed from figure 4.7.3, there is less variation in R and C in recent years, one could assume constant values of R and C in contemporary forecasts.

Figure 4.7.2 - Sum of Squares of Residuals from fitting the Gompertz Survival Function



with time

Figure 4.7.3 - Variation in R and C parameters of the Gompertz Survival Function



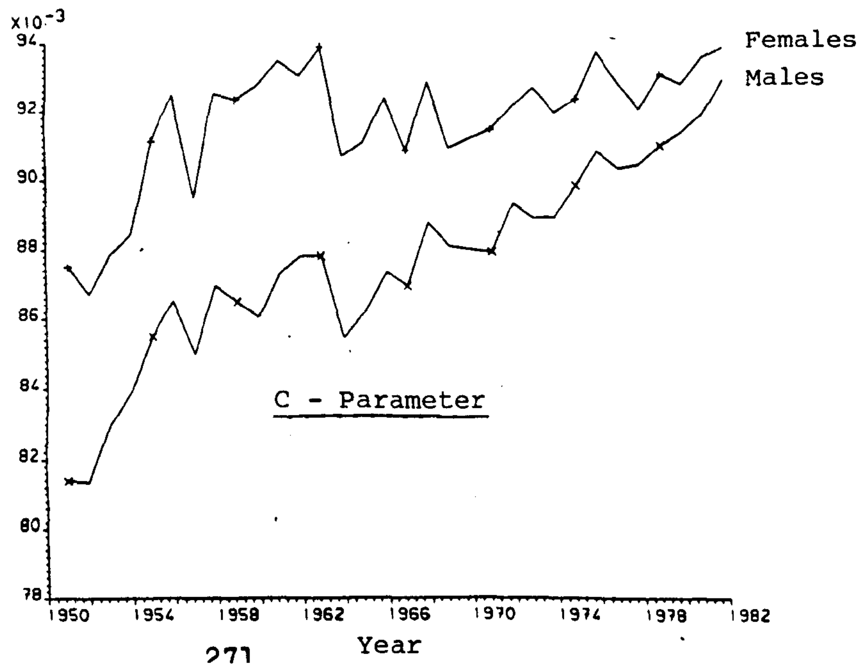


Table 4.7.1 Time based models of C and R parameters

	Period	Parameter	Model
Females	1951-1971	C	$C_{tf} = -1.291 + 1.380e^{0.00013t*}$
		R	$R_{tf} = 0.0005 + 0.0005e^{-0.262t*}$
	1951-1982	C	$C_{tf} = 0.0925 - 0.0062e^{-0.262t*}$
		R	$R_{tf} = 0.0005 + 0.0005e^{-0.207t*}$
Males	1951-1971	C	$C_{tm} = 0.0878 - 0.0068e^{-0.205t*}$
		R	$R_{tf} = 0.0013 + 0.0011e^{-0.251t*}$
	1951-1982	C	$C_{tm} = 0.102 - 0.0188e^{-0.0202t*}$
		R	$R_{tm} = 0.0007 + 0.0014e^{-0.0662t*}$

$t* = \text{year} - 1951$

From extrapolations of S_{at} future values of $P_t(a)$ can be predicted, these can be reconverted back to $n^M a$ giving a forecast of age specific mortality rate.

When the model of S_{at} is combined with information on the current size of the age groups the future size of age groups can be forecasted.

If forecasting in five year intervals the procedure:-

$$N_{0,t,s} = B_{t/SR}$$
$$_4N_{5,t,s} = (B_{t-1}P_t(1) + B_{t-2}P_t(2) + B_t \quad P_t(3) + B_t \quad P_t(4)_{/SR}$$
$$_5N_{a,t,s} = {}_5N_{a-5,t-5,s}(P_t(a) + p_t(a-5))_{/2},$$

gives approximate values of the future age structure,

Where $_nN_{a,t,s}$ is the number in the age group of size n with end age a, at the year t of sex S.

S = gender

and SR = 1 + ratio of females to male births if S = F

= 1 + ratio of male to female births if S = M

The inputs required are $P_t(a)$, SR, B_t and $_nN_{a-5,t}$.

By carrying this out forecasts of the English and Welsh age structure have been obtained.

The forecasts of age structure can be given high and low variants by inputing the 90% confidence limits on births into the above algorithm. 1981 based forecasts, of the expected numbers in the age groups, are displayed in table 4.7.2. and figure 4.7.4.

To show the effect of changes in age structure in a population, whose births are derived from a logistic model fitted to total population, forecasts have been made since 1931 of the age structure of England and Wales. These are compared to observed age structures in figure 4.7.5 (grouping age structures into three age groups). No modelling of $P_t(a)$ has been made for these forecasts, $P_t(a)$ was estimated by interpolation between $P_{1931}(a)$ and $P_{1981}(a)$. Often in practice when forecasting, future survival distributions are assumed and it is not unreasonable to anticipate that a forecaster in 1931 would judge that P(a) will become more favourable, especially when life expectancy is forecasted to rise.

Forecasts shown in figure 4.7.5 for the 20-64 age group show appreciable error as one might expect as no account was made for migration and there was considerable immigration into England and Wales in the post war period. Otherwise the forecasts appear to be satisfactory.

The forecasts of age structure are added to give the total population with high and low variants. The variants are wider than the confidence intervals of forecasts of population obtained in chapter 3. The total population found by summation of forecasted age structures with different birth forecasts for 1931 based forecasts are illustrated in figure 4.7.6. and for 1981 based forecasts in figure 4.7.7.

Table 4.7.2 - Age Distribution for England and Wales 1986-2031

AGE DISTRIBUTION FOR THE YEAR 1986.

AGE INT.	FEMALES	MALES
0	0.311	0.330
1	1.572	1.665
5	1.461	1.538
10	1.555	1.641
15	1.880	1.977
20	1.987	2.069
25	1.800	1.847
30	1.657	1.684
35	1.816	1.831
40	1.580	1.602
45	1.392	1.409
50	1.339	1.347
55	1.396	1.351
60	1.452	1.335
65	1.325	1.110
70	1.249	0.933
75	1.068	0.669
80	0.732	0.357
85	0.379	0.133

AGE DISTRIBUTION FOR THE YEAR 1991.

AGE INT.	FEMALES	MALES
0	0.326	0.346
1	1.637	1.734
5	1.569	1.661
10	1.460	1.536
15	1.554	1.638
20	1.876	1.966
25	1.983	2.059
30	1.796	1.838
35	1.652	1.676
40	1.806	1.817
45	1.566	1.581
50	1.371	1.376
55	1.306	1.292
60	1.344	1.258
65	1.368	1.187
70	1.205	0.921
75	1.068	0.694
80	0.815	0.422
85	0.460	0.182

AGE DISTRIBUTION FOR THE YEAR 2001.

AGE INT.	FEMALES	MALES
0	0.345	0.367
1	1.723	1.824
5	1.683	1.781
10	1.632	1.727
15	1.566	1.656
20	1.455	1.525
25	1.548	1.622
30	1.868	1.949
35	1.972	2.040
40	1.781	1.816
45	1.628	1.642
50	1.762	1.752
55	1.504	1.482
60	1.287	1.229
65	1.185	1.070
70	1.152	0.928
75	1.065	0.732
80	0.787	0.432
85	0.512	0.223

AGE DISTRIBUTION FOR THE YEAR 2011.

AGE INT.	FEMALES	MALES
0	0.356	0.378
1	1.772	1.876
5	1.747	1.849
10	1.717	1.817
15	1.679	1.775
20	1.627	1.715
25	1.560	1.639
30	1.449	1.511
35	1.539	1.606
40	1.852	1.924
45	1.944	1.998
50	1.738	1.750
55	1.563	1.539
60	1.654	1.565
65	1.365	1.227
70	1.103	0.907
75	0.922	0.660
80	0.752	0.435
85	0.510	0.235

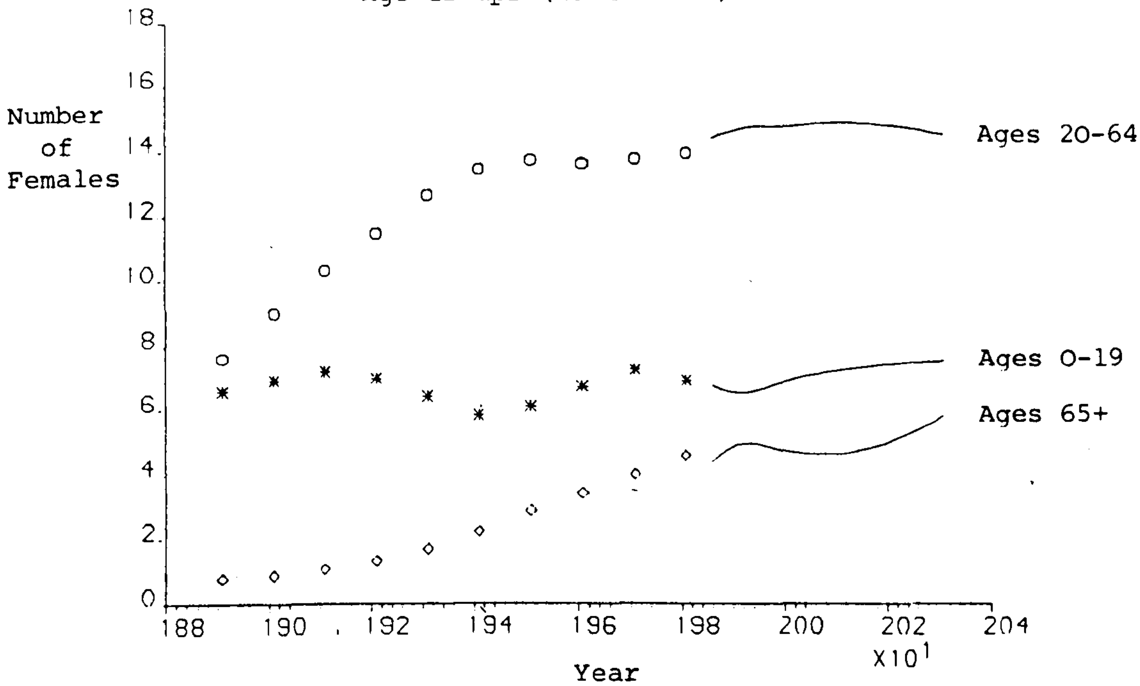
AGE DISTRIBUTION FOR THE YEAR 2021.

AGE INT.	FEMALES	MALES
0	0.363	0.385
1	1.801	1.907
5	1.784	1.889
10	1.766	1.869
15	1.743	1.843
20	1.712	1.804
25	1.672	1.758
30	1.620	1.699
35	1.551	1.624
40	1.436	1.493
45	1.517	1.574
50	1.807	1.855
55	1.867	1.872
60	1.631	1.564
65	1.419	1.274
70	1.418	1.154
75	1.062	0.757
80	0.721	0.426
85	0.442	0.212

AGE DISTRIBUTION FOR THE YEAR 2031.

AGE INT.	FEMALES	MALES
0	0.367	0.389
1	1.825	1.935
5	1.808	1.914
10	1.796	1.900
15	1.781	1.883
20	1.761	1.856
25	1.737	1.826
30	1.705	1.789
35	1.664	1.742
40	1.607	1.679
45	1.530	1.593
50	1.404	1.443
55	1.461	1.482
60	1.707	1.675
65	1.714	1.588
70	1.432	1.209
75	1.148	0.847
80	0.992	0.609
85	0.573	0.289

Figure 4.7.4 - Forecasts of English and Welsh Age Groups (1981 based)



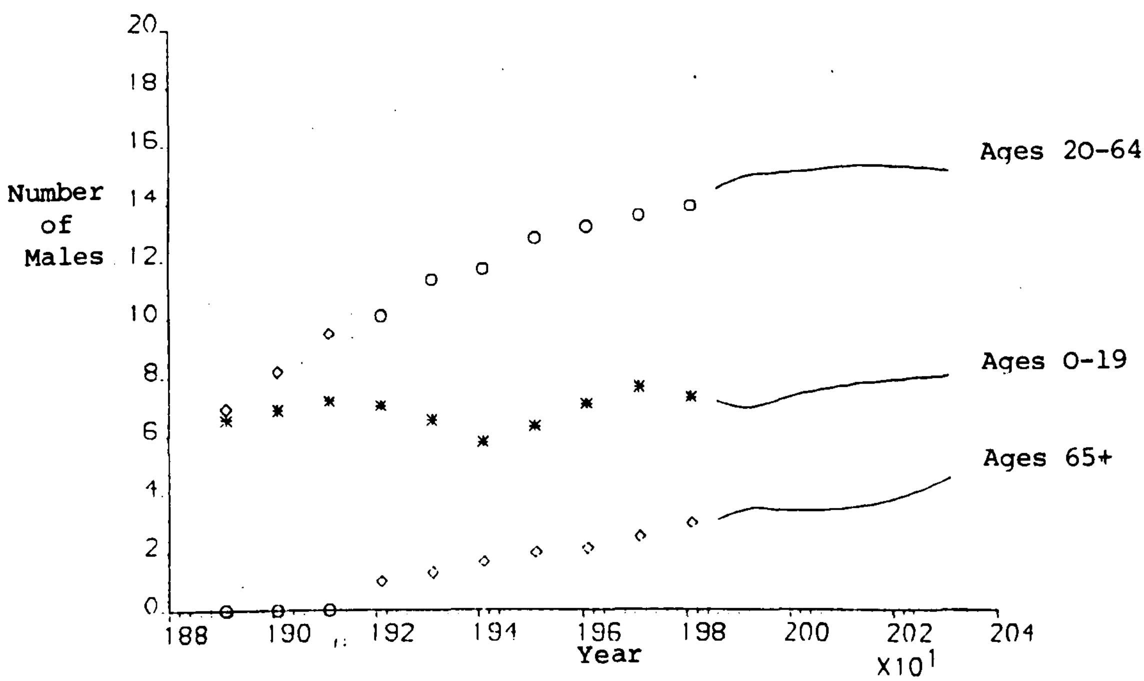
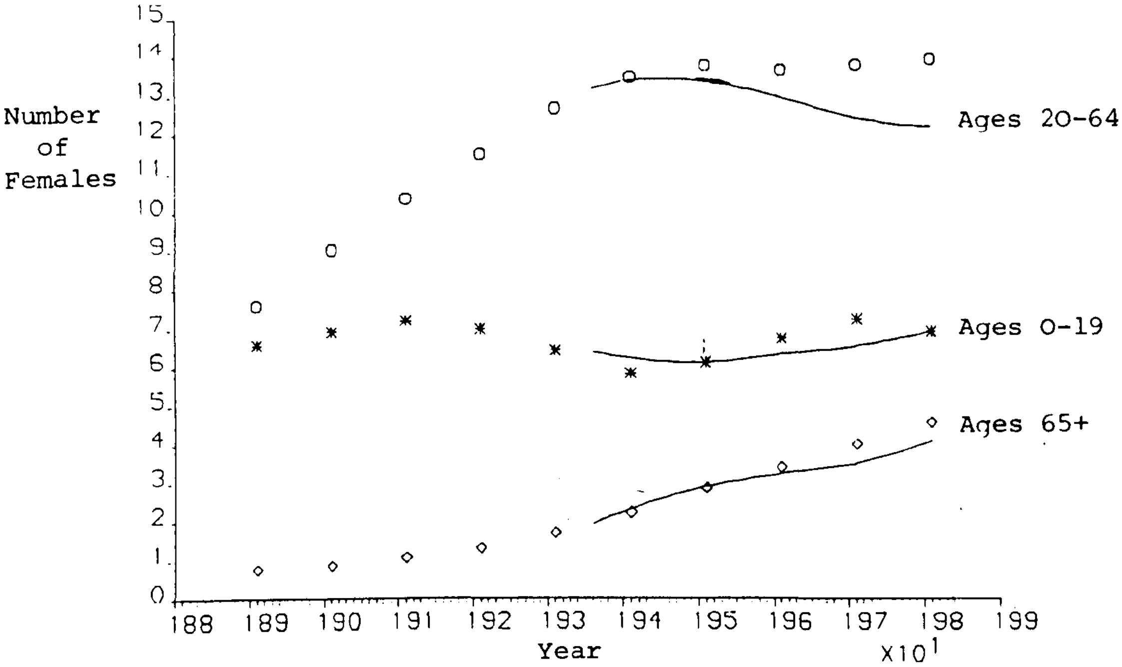


Figure 4.7.5 - Forecasts of English and Welsh Age Groups (1931 based)



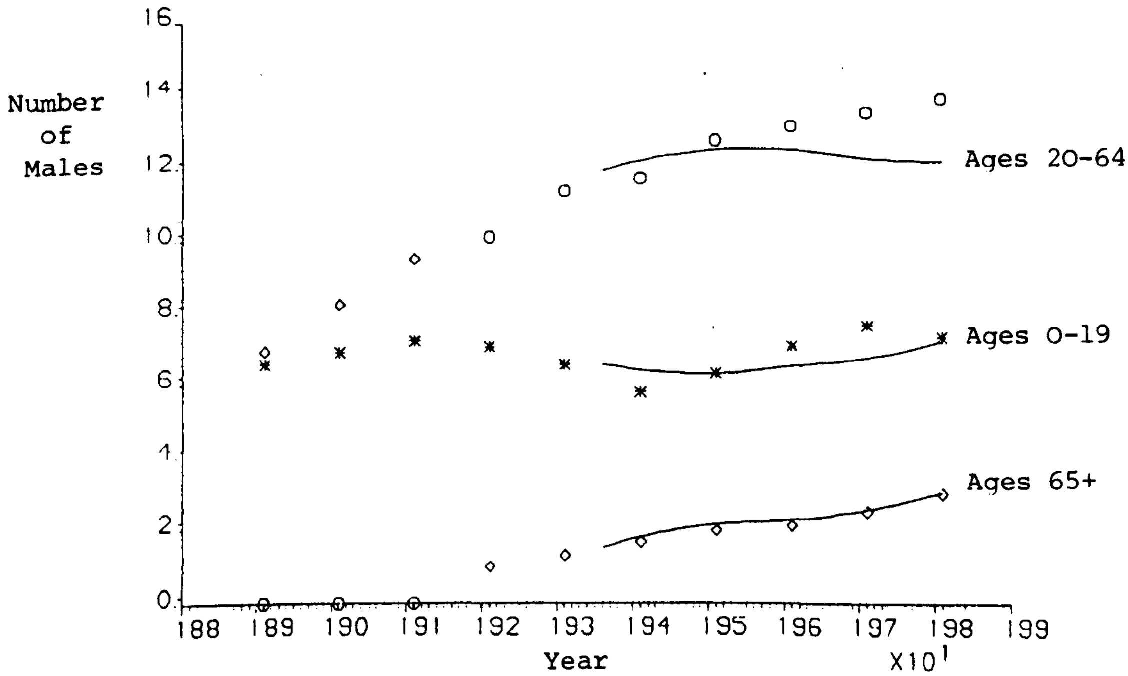
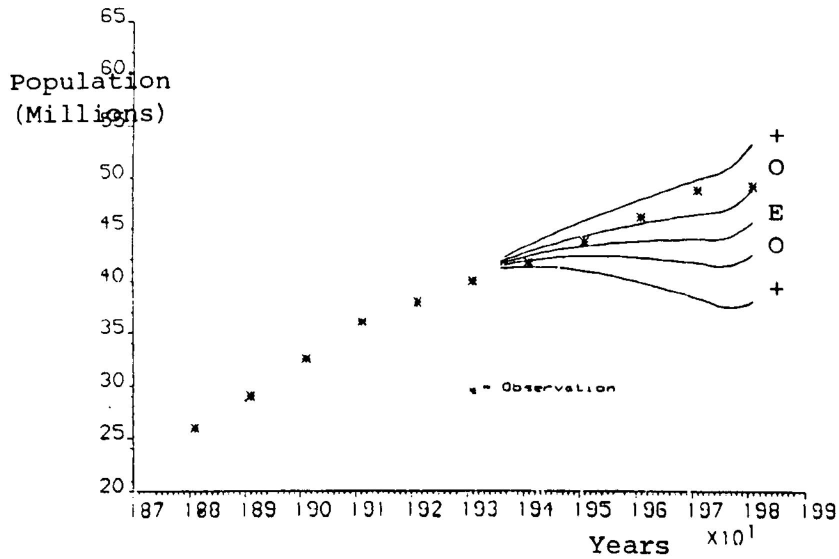


Figure 4.7.6 - 1931 Based Forecasts of English and Welsh Population

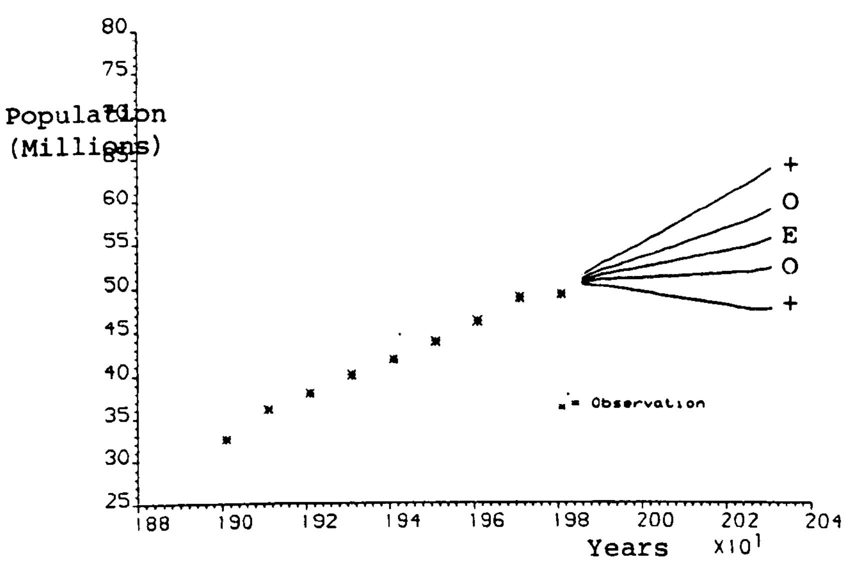


+ = 90% Confidence Limit

O = 50% Conifdence Limit

E = Expected Level

Figure 4.7.7 - 1981 Based Forecasts of English and Welsh Population



For developing nations, there is insufficient data to develop models of P(a), particularly since rises in P(a) might be rapid with increasing adoption of western medicines, social services and improved agricultural technology. In such cases judgements can be made as to the likely future of P(a). One possibility is to assume that the developing nation's age structure corresponds to the age structure of a historical life table of a now developed nation and assume progress to a more recent developed life table, interpolating P(a) for the intervening years.

If forecasts of age structure suggest a quasi-stationary structure then it is permissable to use the method discussed in section 4.6 to forecast the number of births by age of mother, which can be converted to age specific rates by dividing by number in the appropriate age group.

If the age structure is not stationary then one could project forward a frequency distribution, which was fitted to past age specific rates. This would give the probability (b(a)) of a female of age a giving birth. (Shryock et al 1976, p538 discusses the use of Pearson type 1 distributions for modelling b(a)).

In this section when the parameters of S_a were extrapolated into the future an univariate approach is followed in that the parameters are projected by independent models, a better scheme is likely to be to model and forecast the parameters jointly using a recursive updating procedure. The use of Kalman filters (see Harrison and Stevens, 1976) is suggested. (This also applies to projecting the Weibull distribution in section 4.6). In the example discussed the adoption of Kalman filtering is unlikely to be of significant improvement since the age structure is nearing stationarity and the parameters of S_a are tending towards constant states.

4.8 Conclusions and Discussion

It has been shown that if a simple model of total population can be judged as a valid means of modelling and forecasting the total population of a nation, then this model can be used to infer the likely levels of the vital components. Thus, in this respect, the "top down" approach need not be seen as inferior to the "bottom up" methodology.

Possibly a more suitable approach is to combine "top down" and "bottom up" methodology into the one unified approach. This has been done to an extent in this chapter by "feeding in" historical information as to the number of births per year which in conjunction with a deterministic model related directly to the level of total population, forecasts of births per year, with high and low confidence intervals can be obtained. These forecasts of births can then be used to underpin forecasts of age structure. To forecast age structure also requires the bottom level input of information on survival probabilities and current age structure. (Forecasts of age specific births and death rates can also be obtained). In obtaining a model of births, life expectancy at birth had to be modelled. For this sigmoid models were found to be good descriptors.

Figure 4.8.1 illustrates a forecasting procedure for developed populations and highlights the necessary input information.

As can be observed substantial data input is required and the approach is very much of a period based type.

For developing nations the aspirations of a forecaster using this method must be greatly reduced as there is insufficient data to allow satisfactory time series modelling to be carried out. However, the approach is still useful and figure 4.8.2 illustrates a modified approach for developing nations.

Figure 4.8.1 - Forecasting Approach for developed nations

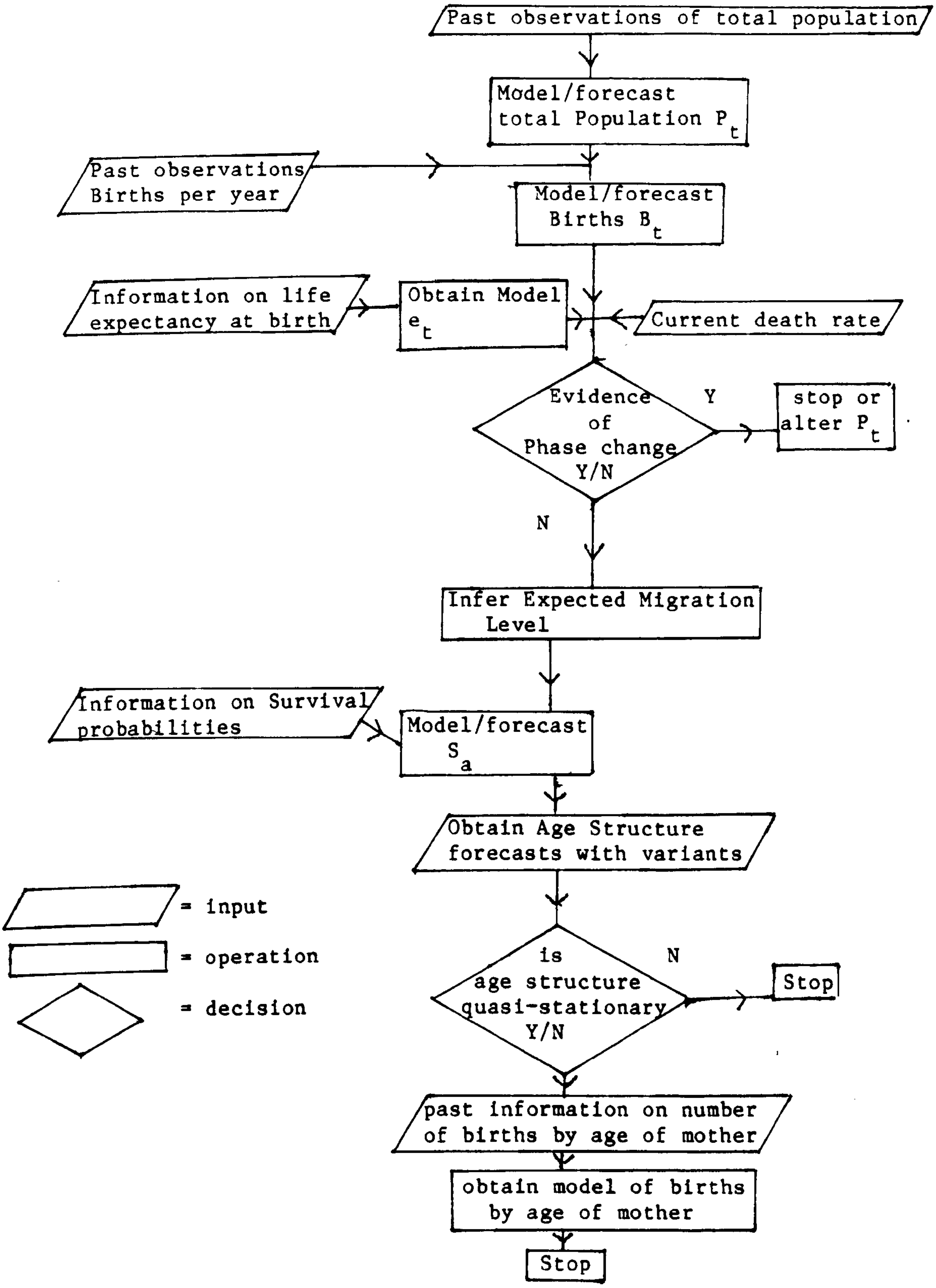
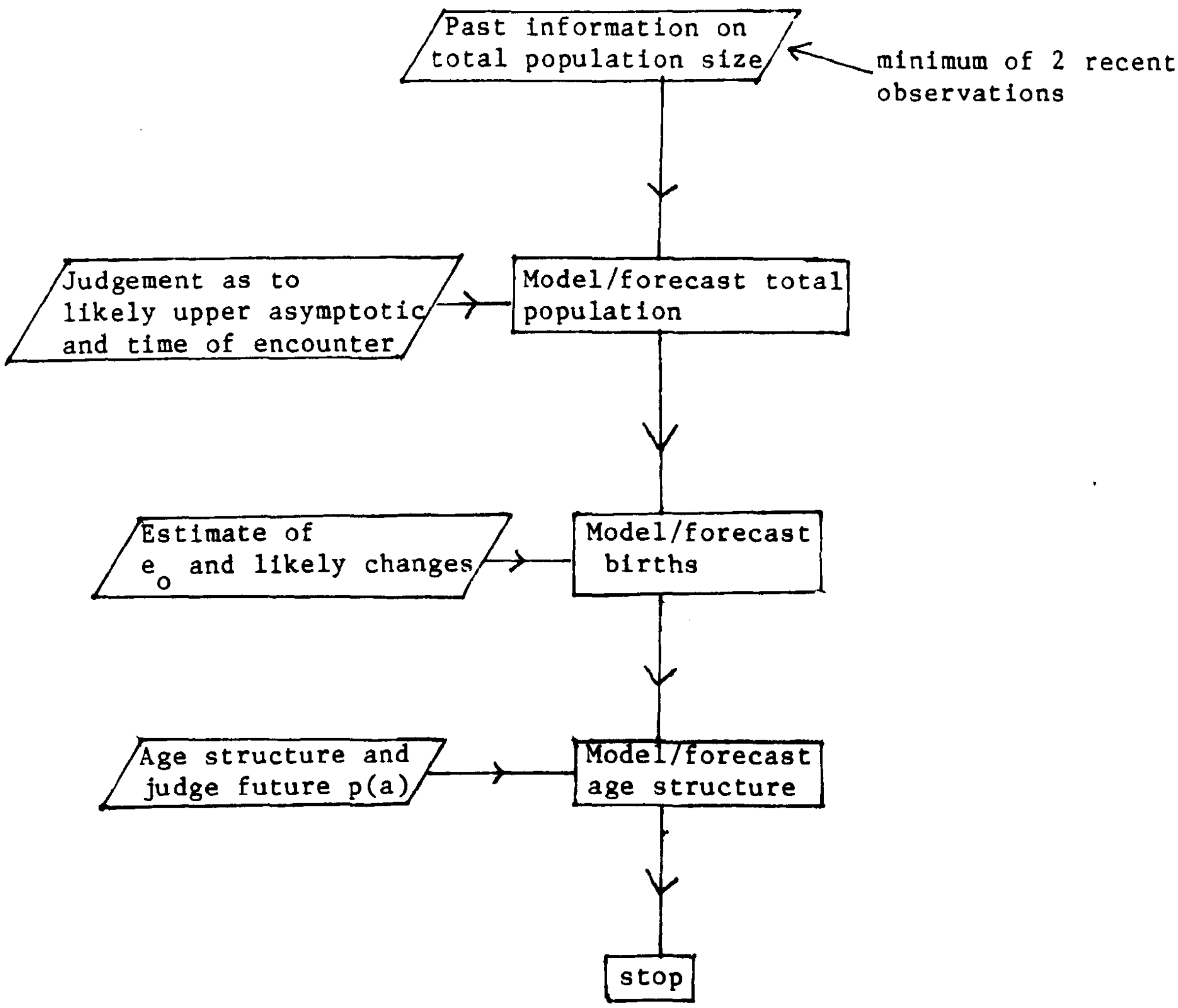


Figure 4.8.2 - Forecasting Approach for Developing Nations



Adoption of these approaches offer advantages over conventional methods as they:-

1. Can be quickly updated.

2. Forecasts can be equipped with confidence limits.

3. Forecasts can be made with minimum of data.

4. Judgements and assumptions are reduced in number and type. Mainly being those of assuming correct model. However, when data is limited judgement must play a greater role.

5. The methodology is to an extent self checking in that forecasts of vital rates can be used to diagnose phase changes and summing of the age structure forecasts made at the end of procedure should verify the forecast of total population made at the beginning.

However to give this methodology credibility the underlying feature of using sigmoids must be related to demographic theory on population change. This is considered in the next chapter.

Chapter 5

The Rationality of Using Sigmoids to Model and Forecast Population

5.1 Introduction

The use of sigmoid type models implies that the long run size of the population approaches an asymptotic level - thus a homeostatic equilibrium. If there is such an equilibrium, then it must be maintained by feedback mechanisms. These, if they exist, can be thought of as two types: positive mechanisms which, to define a terminology, are promoters of growth, and negative mechanisms which are inhibitors of growth or promoters of decline. The mechanisms, it is implied, are influenced by the size of the population in relation to the equilibria.

Population growth (or decline) would be set in motion by changing the relative weights of the feedback mechanisms which might occur through processes endogenous to population growth or by exogenous factors such as climatic changes or epidemics. These changes alter the level of the long-run equilibria.

As populations grow the relative weights between the two types of feedback mechanism alter. Initially the promoters and inhibitors balance each other; population growth occurs when the strength of the promoters increases relative to the inhibitors. Up to the point of inflection the "balance" is very much in favour of the promoters. After this point the strength of the promoters weakens while that of the inhibitors grows, and this returns the population to equilibrium (although at a higher level), when the two are once again in balance.

The rationality of such a process is examined in this Chapter and there is discussion of the possibility of giving sigmoids a theoretical justification. Specifically this quest for a rational base involves an attempt to relate sigmoids to conventional demographic theory, particularly that of Malthus and to identify the promoting and inhibiting mechanisms. To carry this out, the relationship between sigmoids and Malthusian theory is investigated in Section 5.2.

In Section 5.3 the concepts of population pressure and carrying capacity are reviewed and dismissed as explanations of population growth. In the next Section the evolution of more subtle feedback mechanisms are investigated and related to sigmoid models. It is shown that the identification of such mechanisms must be made within the particular historical, social and cultural context of the population. Finally, the demographic validity and interpretation of sigmoid parameters are discussed.

As is illustrated above, the aim in this Chapter is to go beyond empirical validity for using sigmoid models. However, it is not the aim to establish the use of sigmoids as a simple, universal, rational, scientific law of population growth as the earlier population modellers attempted (notably Pearl, 1925). It is of no practical value to establish such a law which would inevitably be intellectually flawed as Wolfe (1927) demonstrated when criticising Pearl.

5.2 Malthusian and Related Theories and Sigmoids

Malthus' view of population growth which is developed in his essay (see Flew, 1982, Spengler, 1972, Bonar, 1966, and Glass, 1953) can be briefly summarised as a system constrained by a "negative feedback" process. To Malthus a population in an unconstrained environment (such as eighteenth century USA) will initially grow quickly (in a geometric manner) spurred by the innate "passion between the sexes". When standards of living fall as resources (principally agricultural land and output) cannot be expanded as quickly as population, population growth will slow and cease. Failure to do so will result in "overshooting" the equilibria. The consequences of this are severe as manifested by starvation, malnutrition, infanticide, disease, war and such like ("misery and vice"). These become more apparent as the equilibria are approached. Malthus did not advocate that catastrophe was the likely outcome, but he did suggest that it was a background threat.

In his latter essays Malthus developed two mechanisms which can operate to maintain populations in harmony with resources. These are the:

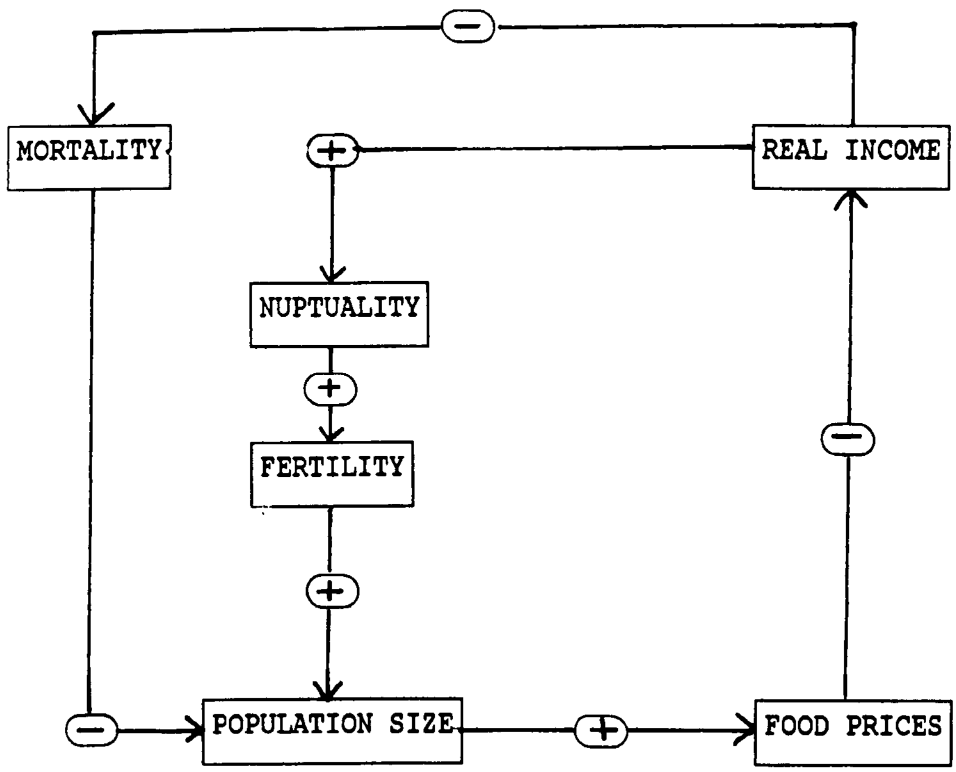
- positive check - which is characterised by increased mortality, sterility and other vices and miseries. The severity of this check can be alleviated by emigration.
- preventative check - families limit the number of children chiefly by delaying marriage (from a contemporary viewpoint it is argued that abstention and contraception should also be included). Malthus argues that this check becomes more important with greater availability of manufactured output which is substituted for children.

Malthus perceived these checks as functioning via the economic system. As Malthus is mainly concerned with population in an agrarian setting he argued that, since the supply of land is fixed, as populations increase there comes a point when diminishing returns set in. With increased population, food prices and land costs rise and real incomes fall. Eventually, if this is continued, the population suffers physically and economically. This can lead to changes in nuptuality as people chose not to marry or to marry later as they could not afford the cost of a dowry, land and the costs involved with setting up a new home and having children. Both the above checks then can operate simultaneously to act to reduce or halt, or even reverse, population growth.

A fall in population would ease demand on agricultural outputs, causing prices to fall and in turn spur a resumption of population growth. Therefore, to Malthus, the rate of population growth is directly related to the general level of material welfare which, in turn, is inversely related to population size, relative to the level of technology and economic growth. This type of feedback loop is illustrated in Figure 5.2.1.

Wrigley (1969), recasting Malthus in contemporary language, points out that, depending on the relative weight of the operation of these checks, the population can settle on either a high stress or a low stress equilibria. The high stress equilibria is associated with high population living close to subsistence which encounter frequent crises. The crises have, in the past, resulted from bad seasons, higher birth rates, or from a period of low mortality. Alternatively, populations with low stress equilibria have higher standards of living due to a strong preventative check which reduces fertility before

Figure 5.2.1: The "Malthusian" negative feedback loop



source: Wrigley and Scofield (1981)

subsistence is approached. There is likely to be a considerable "buffer" between the population's living standard and that of misery; there is "slack" in the system to allow occasional crises to be weathered. In setting the strength of the preventative check, women's roles, education, community care and religious values are important factors, along with economic variables.

However, if the preventative check is too strong, then economies of scale such as access to mass markets, feasibility of certain technologies (e.g. irrigation schemes) and specialisation of skills might be prevented, forcing the population to settle at a lower standard of living. These points are illustrated in Figure 5.2.2.

Malthus suggests that the preventative check develops as the population desires to maintain its standard of living but also can be strengthened by 'moral restraint' arising out of the wish to be able to purchase manufactured output. This concept can be continued to include recent theories such as the New Home Economics theories where, crudely, children are traded off against leisure time, consumer durables and female participation in the workforce. Such an interpretation of Malthus implies that population growth follows a sigmoid path. The relative strength of the two checks influences the position of the upper asymptote in that, the stronger the preventative check, the further the saturation level is from the biological level set by carrying capacity. If the preventative check is absent, then the most suitable empirical model of the bulk of the history of population growth will be an exponential. Minami (1961) (and later Samuelson, 1966) takes the argument further, and explicitly states the relationship between Malthus and the logistic model, by relating

Figure 5.2.2a - Fertility, Morality and Real Income

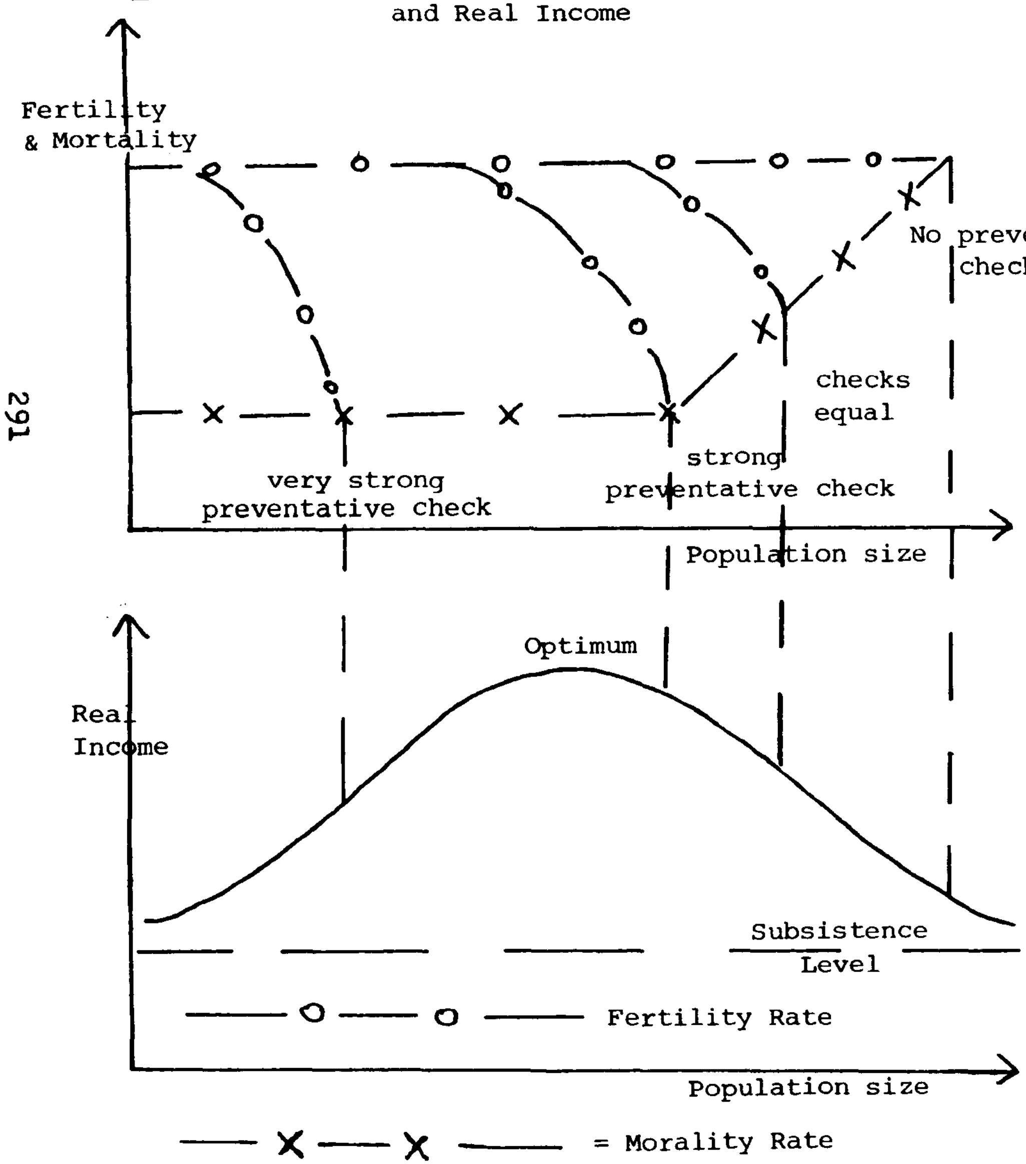
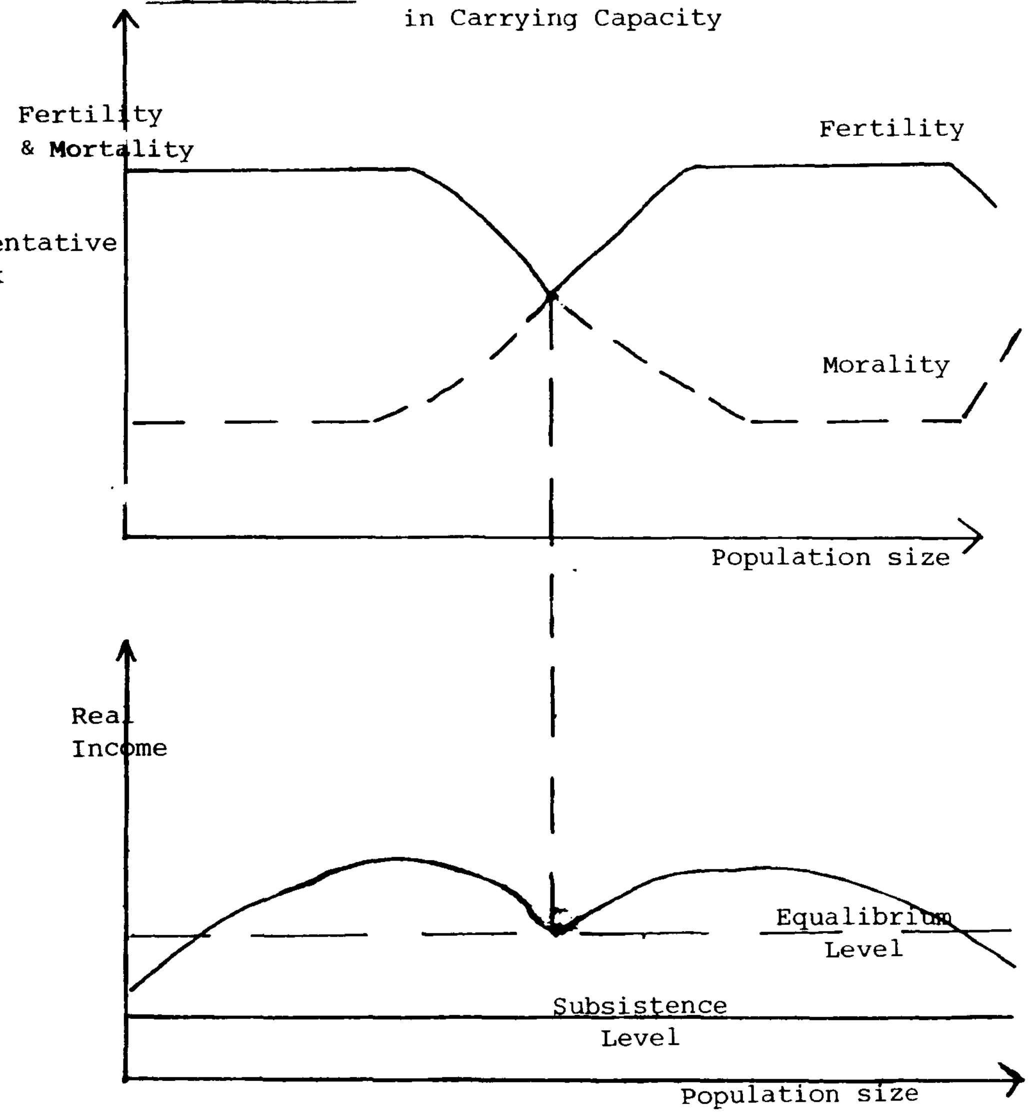


Figure 5.2.2b - The Effect of an Increase in Carrying Capacity



per capita welfare, Y, as measured by real incomes to population, P. That is, Minami writes,

$$Y = b - aP, \qquad a,b > 0$$

and the rate of change of population depends on the difference between the population and its equilibrium level $\bar{P}$ which is directly proportional to the difference between Y and its subsistence level $\bar{Y}$, i.e.

$$\frac{\dot{P}}{P} = c(Y - \bar{Y}) = ac(P - \bar{P})$$

which is a differential equation of the logistic model. (This implies that in this case welfare follows an inverse logistic curve). The relationship is illustrated in Figure 5.2.3.

As homeostasis is approached it is to be expected that there will be oscillation of the population round the equilibrium level. Lee (1974) shows that such behaviour is consistent with a logistic process. Writing a differential equation for population growth as:

$$\frac{\dot{P}}{P} = \alpha - \beta P \text{ where } \alpha \text{ and } \beta \text{ are constants,}$$

when equilibrium is reached, then Lee forms the difference equation:

$$\frac{\dot{B}_t}{B_t} = (a - 1) - bB_{t-1} \text{ (a and b are proportional to } \alpha \text{ and } \beta\text{)}$$

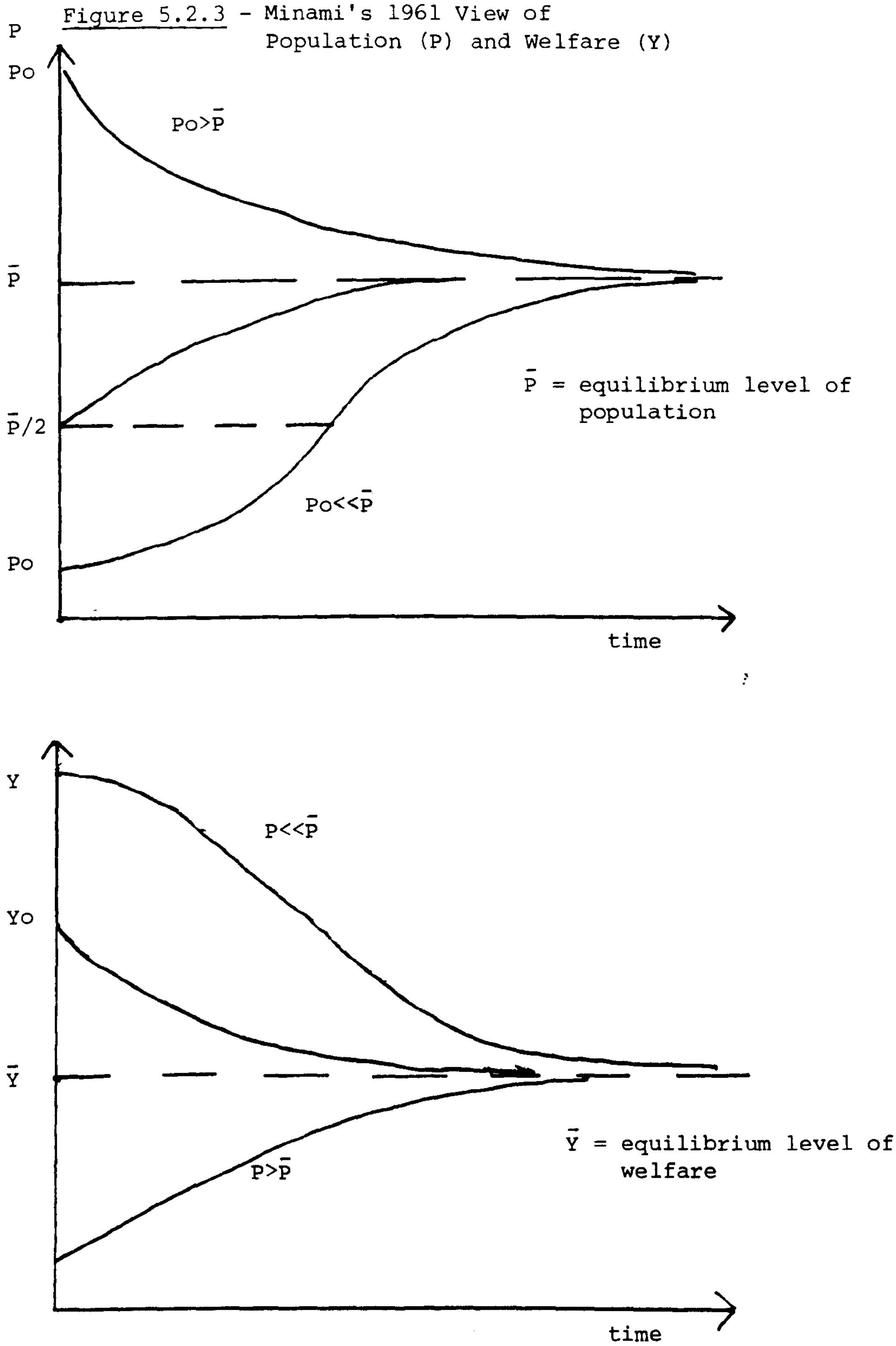
to relate births to the equilibrium population; "a" is the maximum net reproductive rate. This, Lee (1974) demonstrates, leads to two likely possibilities which are:

(a) direct convergence to the logistic curve when the value of "a" is between 1 and 2;

(b) oscillating - when the value of "a" is between 2 and 3 (this might be associated with the Easterlin model which is discussed in Section 5.4).

(For "a" < 1 the population becomes extinct and for

Figure 5.2.3 - Minami's 1961 View of Population (P) and Welfare (Y)



"a" $\geqslant$ 3 the population either falls into a limit cycle or becomes extinct.)

Advances in technology can temporarily release the checks and raise the equilibrium and real incomes. As a result of this a new (higher) normative standard of living is defined (Lee, 1986). This is the "invention pull" view of population growth (Simon, 1981). Phase changes of population growth are therefore seen as a consequence of the expansion of the resource base as a result of autonomously occurring inventions.

Lee (1986) makes Malthusian theory more applicable to industrial societies. Lee writes if a population is to invest in capital goods or more broadly in technology, a surplus income over the level required for subsistence has to be available for extraction by the individual or the nation, in the form of savings, rent or taxes. Such institutional arrangements give greater weight to the preventative check and so prevent the population approaching the maximum possible sustainable level.

Hence, if there is to be "invention pull" expansion, the preventative check must operate effectively. Lee suggests that as surplus extraction is increased, initially the population will work harder, reduce leisure and consume less, but, in the long run, they will have smaller families as the preventative check is strengthened.

However, Lee (1986) also points out that the size of the equilibrium population also depends on how the surplus extracted is used. For example, if the surplus is used to invest in productive technologies, the asymptotic level of the population can be higher than if the surplus is used to "bathe in milk or for foreign tribute".

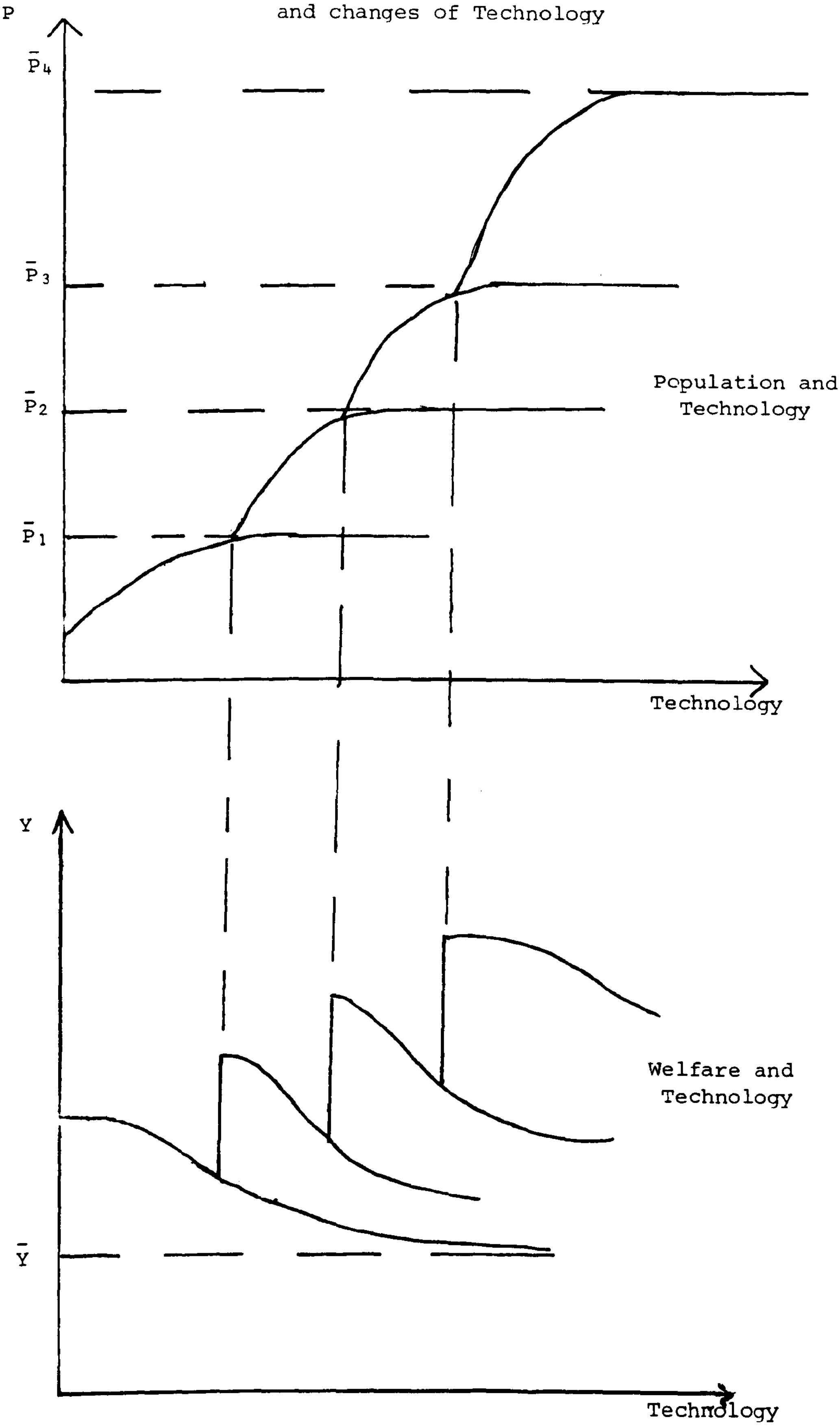
This describes endogenous influences on the population equilibria but there are also exogenous factors such as climatic change or new diseases. Hence Lee (1986) states that the level of equilibrium is "subject to many influences beyond those formally incorporated in the analysis, and is certainly not expected to remain fixed in relation to population and technology".

Minami (1961) describes how technological change can raise the equilibrium welfare level and that this results in a succession of escalated logistics (see Figure 5.2.4).

Thus Malthusian theory gives a description of population growth that supports the use of sigmoids. Within this theory it is easy to demonstrate the need for escalated models as changes in the system result in the surpassing of previous equilibria.

But equating sigmoids to Malthusian theory is not sufficient to validate rationally the use of such mathematical models. The relevance of Malthusian theory to population growth must be demonstrated. Clearly an important concept is that of the equilibrium which, in the absence of the preventative check, operates nearer to the carrying capacity of the environment (i.e. the amount of people that can be sustained in a given area). This discussion implies that the population must be aware of the approach of the equilibria in order to change their behaviour. As the equilibrium is approached the build-up of stress is loosely measured as population pressure. The usefulness of these terms is now discussed.

Figure 5.2.4 - Changes of Population (P) and Welfare (Y) and changes of Technology



5.3 Population Pressure and Carrying Capacity as Explanations of Population Growth

Early researchers who used sigmoid models, such as Verhulst (1838), Pearl (1925, 1930), Pearl and Reed (1920, 1923), Lotka (1925 and 1931) and Yule (1925), related population growth to the density of the population. The carrying capacity which was found by trial and error was defined as the maximum number of people that can be supported per unit area. As this maximum was approached, increased density acted as a inhibiting force - similar to Quetelet's (1835) simple model that population growth is related to density analogous to the motion of a body in a resistive medium. That is, population growth tends to be geometric and is resisted by a force which increases as the square of the speed of growth.

Pearl (1925) tried to use experiments with yeast and animals such as hens, rats, but more infamously, the fruit fly, Drosophila melanogaster, in constrained environments to validate his "logistic law of population". He found that, for example, fruit flies' birth rates fell and death rates rose as the density of the flies in a bottle increased. The growth of the bottle population of flies ceased when the two rates balanced. When the bottle was expanded (i.e. constraints on the environment slackened), growth resumed.

Pearl (1925) draws further support for his model by generalising from individual growth processes such as the growth of tadpoles' tails, weights of plants and humans. (Other animal researchers are discussed by Lloyd (1967) and listed by Hart (1945)).

Lack (1954, 1966) finds density dependent mechanisms to be important regulatory mechanisms affecting animal populations (particularly birds) Lack found these mechanisms to operate via mortality in that higher densities lead to food shortages, increased preditation and disease and also via fecundity (more so for mammals such as deer and insects than birds, although for gregarious birds high densities result in crowding, causing the territory system to collapse so disrupting reproduction). Bulmer (1975) found evidence of density dependence in Canadian fur-bearing animals (coloured fox, arctic fox, lynx, mink and muskrat), and verified this statistically by using a test based on the correlations within the series. Fujita and Utida (1953) and Gill (1972) also present evidence of density dependent growth in azuki bean weevils and paramecium, respectively.

As populations other than human were found to grow in a way described by a sigmoid curve and density dependent mechanisms, early researchers thought there could be some factor operating at a chemical-physiological level (e.g. von Bertalanffy (1938, 1957) relating the process to autocatalysis), which is still a view for some small mammals. The early researchers recognised that the density dependent mechanisms operated in more subtle ways than in animal populations as there were abundant examples of population growth slowing without any apparent sign of resource shortages (assuming that resources were efficiently exploited). Thus higher order factors were suggested involving social-psychological concepts such as congestion of living space and loss of privacy. This is a view advanced by Douglas (1966). When discussing population control in primitive groups Douglas argues that it is wrong to relate the population equilibria to carrying capacity as the optimum population, as Carr-Saunders (1922) and Wynne-Edwards

(1962) do. Douglas advances the notion that optimal population is defined by the desires of the particular population; these desires depend on the degree of development of the population and the harshness of the environment. The argument is, the greater the degree of development, the more regulation will be carried out and the further the equilibria moves below the biological carrying capacity. Douglas states: "Human groups do make attempts to control their populations, often successful attempts. But they are more often inspired by concern for scarce social resources, for objects giving status and prestige, than by concern for dwindling basic resources". (This might partially explain why in some African countries, population pressure is seen as contributing to erosion, deforestation and desertification, yet the population encourages more recruits.) Douglas reckons that, as populations develop, newer higher levels of desires develop which define the limits to population growth. Thus it is "the demand for oysters and champagne, not for the basic bread and butter, that triggers off social conventions which hold human populations down" (Douglas, 1966). (Hence there are parallels between psychological motivation theories such as Maslow's (1954) hierarchy of needs and Douglas's thesis.)

Douglas (1966) gives some examples of populations which practise population regulation; these are illustrated in Figure 5.3.1.

Figure 5.3.1: Douglas's Hierarchy of Population Regulation

Nambudire Brahmins of Southern India are a rich elite who preserve their exclusivity by preventing marriage.

The natives of the South Pacific island of Tikapia who practise abortion and infanticide despite abundant root crops, as they perceive coconut cream as a necessary but scarce resource.

Increasing development (and affluence)

Rendille camel herders of Kenya practise infanticide, limited marriage and emigration in order to maintain the population in equilibrium with the stock of camels which is perceived as being fixed.

The Netsilik Eskimoes live in a very harsh environment and practise female infanticide for survival. However, due to high early mortality amongst males, there is sexual balance in adulthood.

Subsistence

Focusing on such density dependent relations might explain one phase of population growth, but not why growth would resume. To account for multiple phases of growth, researchers have argued that analogous to the fruit flies in the bottle, the bottle is expanded and the carrying capacity raised. This can come about by some combination of the following:

(i) Climatic changes - Butzer (1981), Hassan (1981) and Yesner (1980) consider that small changes in the climate can have dramatic changes for primitive agrarian or hunter-gatherer societies in altering their resource base. To a high stress equilibria population this could be advantageous alleviating resource deficiency and allowing population growth. Alternatively, if the climate becomes more severe, this can lower the carrying capacity ceiling so causing population decline. For a low stress society, they could be converted into a high stress population. Climatic change will affect all societies. Gilmore (1981), for example, writes of the "little Ice Age" which was a cold period from medieval times to 1850 and may have caused a lowering of agricultural productivity in England. (Wrigley and Scofield (1981) and Galloway (1986) also found evidence of this.) Broecker et al. (1985) illustrate that climatic change is a subject of serious concern which can occur suddenly and quickly. (Gilmore (1981) presents a Catastrophe Theory model of such changes suggesting that major climatic change can occur in under one hundred years.) There are suggestions that climatic change might arise as a result of human activity raising carbon-dioxide and waste heat emissions (see Austin and Brewer (1971)).

(ii) Technological change - Technological breakthroughs allow the resource base to expand. These can be "invention pull" (as discussed earlier) or "population push" which refers to technological breakthroughs occurring as a result of people trying to improve their situation when population pressure is manifested. Boserup (1981, 1986) believes "population push" is important in agrarian communities and also argues that, if the size of a population is insufficient to support a technology, there can be technological regression and population decline. Technological optimists, such as Simon (1981), argue that populations at any stage of organisaton or development can, as population pressure is perceived, bring about technological change. Problems arise because of adjustment and the lack of population pressure. Thus Boulding (1973) writes: "Perhaps the two most significant phenomena of human history have been, first, the tendency of man to expand into an existing niche and, second, his tendency to expand the niche itself once an old niche has become tight. This has brought us from the paleolithic to the neolithic to civilisation, and now to post-civilisation." Wrigley (1969) also comments: "The challenge of population growth may act as a catalyst of change and development though it is also possible that increased misery may induce apathy." In connection with this, Malthus pointed out that the political regime and political attitudes were important in determining how a population operated.

(iii) Changes in the way society is organised - Changes in the organisation of society can alter the surplus extracted and how it would be invested (see Lee (1986) and Smith (1981)). Carneiro (1970), in discussing the origin of the state, presents a view that small primitive groups can combine to form a state which, in turn, leads to the development of a political system and division of labour, increasing the resources of the population as a whole.

Carneiro's view is that, in primitive societies, the dominant means of social organisation is in villages constrained by the environment, as was the case in Peru (by mountains) and the Nile, Tigris-Euphrates and Indus valleys, or by resource concentration (as with villages on the shores of the Amazon), or by social concentration around administrative or ritual centres. As population pressure is perceived, warfare breaks out and the victors incorporate the defeated into their own territory and, as this continues, states begin to emerge. Alternatively, coalitions might be formed. Renfrew (1978) and Renfrew and Poston (1979) illustrate, using Catastrophe Theory (which is discussed in Chapter 6), how sudden changes in perception as to the advantages of being in larger communities can result from gradual changes in variables such as the distance to travel to farm and the desire for effective defence. This, Renfrew postulates, gave rise to the sudden changes in the organisation of population from village into city and allowed renewed population growth.

(iv) Diversification - Hayden (1981) suggests that Stone Age populations, rather than expand the resource base by developing new technologies, diversified the resources used - i.e. more plentiful r-species animals (small abundant animals with short life cycles) which, in food chain terms, are nearer the primary source of energy, were sought rather than k-species (large animals with long life cycles) (May (1981)). This involved some technological change in the development of cooking and new tools such as grinders, pestles and mortars. This gave a more stable resource base. Hayden (1981) considers that diversification is preferable to other strategies as less effort is required, and assumes populations conform to Zipf's (1949) principle of least effort. The urge to diversify could have resulted from climatic change, making k-species rarer (Yesner (1980)), or from over-hunting of k-species.

(v) Cyclical economic changes - Some, such as Bell (1974), Marchetti (1985) and van Wyk (1985), advance the view that technology develops as a series of escalated sigmoids. Each cycle represents a major technological breakthrough which raises the "carrying capacity". This process is related to Schumptarian (1964) notions of growth and decay, and is illustrated by Piatier (1981) who believes that long-term growth can be thought of as a succession of industrial revolutions. Piatier writes: "An industrial revolution is a secular movement which can be schematised by means of a logistic curve: four ages and their corresponding industries can be read into it. Infancy with weak growth, adolescence with rapid growth, adulthood with slower growth and maturity with zero growth."

It is appealing to relate population growth to such a process; unfortunately there is little evidence to support such movements of technology (see Freeman, 1983) and, in developed populations in recent years, technology has increased at unprecedented rates, yet population growth has slowed.

(vi) Changes in perception - Douglas (1966) explains population growth and moves to new population equilibria as changes in perceptions as to what is necessary for maintenance of a desired lifestyle. She comments: "When social change occurs so rapidly that the prestige structure is no longer consistent, we should expect population explosions to occur. Or, if the whole traditional prestige structure is broken as a result of foreign oppression or economic disaster, again we would expect that the social controls on over-population would be relaxed."

Although arguments such as those above, coupled with the models examined in Section 5.2, would support the use of sigmoid models, when examined further there is little evidence of such views. The approach of focusing on density dependence, population and carrying capacity can be criticised on several counts. The major criticisms are now listed:

(i) The straightforward population pressure argument fails to explain why, despite the apparent excess availability of resources, all populations, even primitive populations, have fertility well below natural maximum fertility levels.

This is highlighted by the phenomena of the demographic transition which has been observed in many Western nations over the last century and involves population growing as the result of falling birth rates lagging a fall in death rates (Carr-Saunders, 1922, and Coale, 1974). This process has generally been accompanied by improvements in living standards. The transition might have started with a rise in fertility (Dyson and Murphy, 1985) and is contrary to explanations of population growth given by the bottle model. The magnitude of the fall in fertility which has put many developed nations below replacement level fertility (e.g. West Germany, Sweden and Great Britain) is highlighted by Calot and Blayo (1982).

Rather than improvements in economic status promoting unchecked populaton growth, it is considered in this thesis that it is likely that there is a process whereby fertility falls to near replacement as economic status improves and that this will eventually apply universally. There is support for this view, for example, Caldwell (1982) argues that underdeveloped and developing countries are inevitably strongly influenced by Western societies in a process he calls Westernisation, and the effect is to cause a fall in fertility.

Little evidence can be found of density related phenomena to explain this fall although some have tried. Galle and Grove (1978), for example, found small but significant correlations after social class was controlled for, between high density living and indicators of pathology such as death rate, infant mortality rate, suicide rate, admissions to mental hospitals, illegitimate births, juvenile delinquency and adult crime, after studying the effect of crowding in residential areas, transport and classroom crowding. However, he failed to control for income effects. Karlin (1980) concludes that "in general, crowding has a complex but predominantly negative effect on the social behaviour of humans".

Also various researchers, such as Galle and Gove (1978) and MacIntosh (1978) have found some detrimental effects arising from overcrowding in animal experiments. However, it is tenuous to extend these findings to human populations and argue that overcrowding results in fertility falls among humans. Other researchers such as Epstein (1980) found no clear-cut effects of density on human populations. Freedman (1980) points out that density alone does not reliably produce negative effects on humans and that "normal human populations appear to be able to function well over several orders of population density". Indeed Freedman (1980) and Epstein (1980) can find no physical, social or psychological manifestation of overcrowding.

(ii) Measures and definitions of population pressure and carrying capacity are dubious. This arises due in part to effects being socially mediated (Douglas, 1966), probably even among animals (Wynne-Edwards (1962) argues that territorial behaviour and hierarchy in animal populations are expressions of this as they act to limit surplus individuals' access to food and to reproduction). Wynne-Edwards (1978) comments that, in order to prevent over-exploitation of resources and to maintain an optimal population, animals engage in "social competition as an artificial regulator of food demand". Carr-Saunders (1922) also argues that there are social pressures on fertility and mortality that act to constrain human populations to an optimum level in relation to resources and that starvation is not an acceptable means of limiting populations as it "makes social conditions unstable, and all the accumulated skills in hunting and cultivating the ground become worthless". Hence, carrying capacity is open to many definitions which are almost impossible to calculate (Hayden, 1975). Also there is the problem of the complex cyclical nature of the resource environment - at which point should the carrying capacity be measured? Thus, how can population pressure be perceived by the population and translated into actions?

The difficulty in measuring carrying capacity cast doubts on research evidence as to the existence of density dependent mechanisms. This is highlighted by Clarke (1973): "The search for an operational index of the population-resource ratio is like trying to describe landscapes in quantitative terms, it reduces complex reality to meaningless units".

However, in some settings population pressure can be identified. Lack (1954) points to the extreme cases of outbreaks of cannibalism among rodents, lions and primitive man, and Grigg (1980) finds evidence of population pressure in nineteenth century

Ireland in rising land prices and food shortages. Mirowski (1976) found evidence of density dependent mortality, finding a significant positive correlation between increases in grain prices and deaths from typhus, smallpox and fever, and total deaths in London from 1670 to 1830.

Wrigley and Scofield (1981), when examining the population of England in the seventeenth to nineteenth centuries, however, found only a weak density dependent relationship manifesting itself in mortality rates and concluded the greatest part of mortality to be exogenously determined by climatic change and disease.

(iii) The "technology push" view of allowing further population growth is flawed, as Cowgill (1975) pointed out. Those who experience resource deprivational stresses are unlikely to be in a position to devote much resources to research and development even if only at a trial and error basis. Innovation and technological development are more likely stimulated by the prospect of new opportunity rather than the threat of hardship. This is illustrated by the Industrial Revolution in Europe, and especially in England, where between 1650 and 1750 the population of England was stationary or even declining (Wrigley and Scofield, 1981), agriculture improved and living standards rose and, out of such a background, industrialisation appeared.

Social stratification is important and elites may push technological innovation which does not help the bulk of the population, such as was the case in China where gunpowder and ceramics were developed.

Cowgill (1975) further pointed out that development is socially defined and is unlikely to occur if the controlling elites in society do not see any advantage in development, and Cowgill writes: "Compassion for the hard lot of the poor is not by itself apt to encourage innovations which can only be undertaken by the prosperous, unless these people also feel threatened by hard times but are not so threatened that they are unable or unwilling to take on new enterprises". Far more likely to stimulate development are good times, or the prospect of good times.

McNicoll (1984) is also dubious about the existence of a link between population and technology, and comments that, although in general the scope of productivity gains have been underestimated, the concept of technological rescue always being at hand is doubtful. Continued high population growth does seem to be detrimental to economic growth.

(iv) Carneiro's (1970) views on social changes and the origin of the state which emerges by warfare is not conducive to growth as surplus extraction may be severe to support a strong military so preventing investment in social and resource infrastructure and constraints may soon be imposed by resources. Also it is not clear why warfare should be a unifying factor in economically poor societies. Cowgill (1975) takes up this point to suggest that increasing warfare is more likely where there is increasing and not decreasing wealth as the gain from conquest would be more worthwhile. (However, it was possible that warfare was endemic in many primitive societies and is not linked to density.)

(v) Hayden's (1981) model of resource diversification leading to early settled agriculture and the apparent population growth in late Paleolithic to the Mesolithic period lacks hard evidence (Pittioni, 1981), but is also criticised on theoretical grounds, for example, by Goodyear (1981) who points out that the costs of sedentary lifestyles are high compared to hunter-gathering as more work may be involved and the population is more vulnerable to disease and pestilence. A severe criticism is made by Maston (1981) who points out that switching to r-species will not necessarily give a more reliable resource base as, unlike k-species, r-species tend to go through "boom and burst" cycles (May, 1981) and so there can be periods of extreme scarcity.

(vi) Repetto and Holmes (1983) question the view that population growth alone has caused resource shortage, pointing out that resource shortages for a population have often occurred as a result of commercial exploitation by another population, poor management of the resources and inequality of access to natural resources. They comment on the present developing nations that "these factors operating in conjunction with population growth are leading to considerably more rapid deterioration than would demographic change alone". They go on to show that "in many developing countries, despite the growth of population, pressure on land has not come mainly from small subsistence farmers but mostly, and in some countries entirely, from the increase in acreage of large commercial holdings". Hence, economic factors, property/resource ownership and access rights could well play a greater role than population growth in determining the welfare of a population.

Thus, one concludes that the explanations based upon density dependence, which would have backed up sigmoid growth models and led one to believe that escalations are to be expected when the resource base is expanded, are not to be found in societies above the most simple level. Density dependence and population pressure are only part of the explanation since there are some density dependence relationships between wages and fertility as Wrigley and Scofield (1981) found in seventeenth and eighteenth century England. However, the relationships are complex and not easily determined. Thus there is agreement with Wolfe (1927) when he stated: "To make a direct correlation between fertility and density is to abbreviate too much. It is to ignore highly important middle terms - the standard of living, social conventions, the state of the arts, industrialisation, the growth of individualism, rationalism, and democracy, and the spread of contraceptive knowledge".

In the next Section, population development is reviewed in the hope that regulatory mechanisms can be found and their operation understood. This will help in considering the relevance of Malthusian theory and, hence, the rationality of sigmoids.

5.4 Mechanisms in Social, Cultural and Historical Context

In this Section some of the factors which have acted as population growth promoters and inhibitors are considered with a view to determining how they relate to a feedback model. The factors which cause population change and regulation are discussed in relation to the development of population. It is assumed that populations tend in the long run to homeostasis and that there have been three main phases or cycles of growth related to the dominant means of production, primitive hunter-gathering and nomadic agriculture, agrarian and industrial (thus corresponding to the cycles illustrated by Deevey, 1960), although it is recognised that not all societies have expanded in this way as some remain in a particular mode for a long period.

The aim of this Section is to show that there are important factors which influence population growth other than those directly related to population size and the ones associated with size often act in subtle ways. It is not the intention in this Section to cover comprehensively all courses of population change and regulations for all populations, but to demonstrate that there are social and cultural differences among these causes, for different populations. It will also be demonstrated that the mechanisms have evolved over time.

5.4.1 Primitive Populations

Such populations are exemplified by Paleolithic man and currently the !kung San of the Kalihari and the traditional Australian Aborigines; they are predominantly nomadic hunter-gatherers. Researchers such as Hayden (1981), Hassan (1973, 1980), Howell (1986) and Lee (1980) have shown that such populations have maintained small or zero growth rates for long periods despite the relative abundance of resources. Apart from Douglas (1966) (whose explanation was outlined in the previous Section), other researchers have studied contemporary hunter-gatherers to try and determine why homeostatic equilibria should have occurred. Smith (1972) argues that there are demographic responses to population pressure in archaeological times. He suggests that population pressure, which is defined when the carrying capacity is at a minimum, advances technology and limits population. (Smith lists much of the literature on carrying capacity as an explanation for population growth at this time.)

Howell (1986), after studying anthropologically the !kung San, emphasises a biological explanation of maintaining a low level equilibria, in that, by the nature of the !kung physiology and diet, they have very low fat levels which inhibit ovulation while the mother is lactating. Lactation, which reduces a mother's fat levels, continues for a prolonged period, partly as a result of harsh diet. This, combined with a post partum sex taboo, keeps long intervals between child spacing, hence preventing rapid population growth. However, Hassan (1973) is doubtful that views such as Howell's can be generalised to early populations as there is no evidence that such populations were badly nourished, and points out that females on high protein diets tend to reach menarche earlier.

Lee (1980) argues that long child spacing intervals are essential for nomadic hunter-gathering populations as infants are carried by mothers, thus even two children present a burden to mothers. Lee (1980) reckons that the child spacing may be imposed as a result of the heavy workload physically weakening mothers' fecundity, and nipple stimulation by the child carried at the mothers' shoulder preventing ovulation, coupled with the post partum sex taboo.

However, such generalisations from the !kung are not representative as they have relatively low mortality rates, rarity of aggression, abortion and infanticide and unusually long birth intervals, which are uncharacteristic of other societies such as the Hadza of Tanzania (Dyson, 1977) and Australian Aborigines (Birdsell, 1978). Also contemporary hunter-gatherer populations live in environmental extremes, probably not typical of past populations.

Birdsell (1968), after examining Aborigine groups, points to warfare and infanticide as important regulatory mechanisms. Referring to the Pleistocene period, Birdsell (1968) estimates that, between 15% and 50% infanticide occurred through ritualised killing and abandonment of offspring (particularly females). The importance of infanticide is stressed by Ripley (1980) who believes its practice is increased as the population approaches the carrying capacity of the environment. This is also supported by Hassan (1980) who estimates the average mortality in prehistoric societies to be that between 40% and 50% die before adulthood. Polgar (1972) considers infanticide to be the main regulatory mechanism in these populations and that mortality crises, brought about by disease or starvation, are not important. (However, high mortality of mothers during giving birth may have been an influence (Hassan, 1973).) Since many parts of the environment were uninhabited so, as food became scarce in one area, groups could split and migrate to an area of more plentiful resources, so food shortages were unlikely to promote regulating behaviour. Also the small size of hunter-gatherer groups prevented the establishment of viral infections transmitted by contact or via the respiratory or dietary tracts.

Hassan (1980) argues that population equilibria with resources was maintained at this time through both long child spacing and infanticide and abortion.

The fortunes of primitive populations could have been closely linked to fluctuations in animal populations. For example, Eskimoes underwent boom-crash growth in relation to Caribou populations which have a 70-year cycle (Burch, 1972). The pastoral Samburu of East Africa, Spencer (1973) illustrates, have a sawtooth growth profile related to cattle numbers which collapse in dry years. This highlights the sensitivity of primitive populations to environmental change.

Changes in climate may have led to the establishment of more sedentary life-ways in the Mesolithic period. Butzer (1981) considers that ice ages made large traditionally hunted animals scarce and forced the population to develop other resources, although over-hunting was likely to be an important factor. Stuart (1986) considers that overexploitation of k-species likely occurred towards the end of the Pleistocene period. He cites as evidence the extinction of many large animals around the world which coincided with intrusion of humans into their environments. Stuart points out that the disappearance of these animals was unlikely to have been solely the result of climatic change, as there is no evidence of mass migration of these animals to more suitable areas, and why should the climatic change at the end of the Pleistocene period and not earlier climatic changes be associated with mass extinctions?

However, Stuart also comments on the lack of evidence to support the "overkill" hypothesis. Possibly the extinctions were a result of both climatic change and human activities. Hayden (1981) argues, however, that the resource base became diversified as it involved less effort than hunting and gathering and increased resource reliability.

This led to the establishment of settled agriculture which Howell (1986) and Lee (1980) argue lifts the biological and physical checks of "fatness" and workload and allows fertility to rise. Hassan (1973, 1980) is doubtful that this had much effect; more important maybe, was the reduction in infanticide. Infanticide was reduced and possibly fertility increased as it became advantageous for populations to increase so that work could be shared and for defensive reasons. Children became a greater asset as they could contribute directly in an agricultural setting. Hassan's (1980) view of population change in the Neolithic period that there was slow fluctuating growth with periods of stagnation and depopulation is likely. Growth occurs, not by shortening birth spacing, but by abandoning infanticide and abortion, and by improved child life expectancy through more care being afforded to children.

Therefore, in primitive populations, mechanisms regulating population growth appear to be defined through a complex system that involves interactions with the environment and economic and social factors. From this it is possible that the above mechanisms promoted population growth and that, in these early societies, population might well have grown in a manner reflected by escalated sigmoid models. This pattern of regulation is only partly explained by relationships to carrying capacity.

5.4.2 Agricultural Populations

In agricultural populations some evidence has been found supporting the theory that increasing population density reduced living standards and slowed or reversed population growth. For example, Wrigley and Scofield (1981) and Lee (1973) found that in late sixteenth century England, as a result of high population density, food prices rose and marriage was delayed. Thus the preventative check was apparent, but there were considerable lags (up to fifty years) in the response to falling living standards. They found little evidence for the operation of the positive check, especially in the short run.

To investigate mechanisms of agricultural population change, first Western European demographic history is considered, and then the situation in current developing nations is commented on.

Grigg (1980) suggests that population growth in Western Europe from 1000 to at least 1900 can be viewed as three cycles, the first from 1000 to 1350, the second from 1450 to 1650, and the third beginning around 1750. These cycles, Grigg illustrates, can be schematically represented by a succession of logistic curves. Factors affecting population have changed as the cycles progress, as have responses of the population.

Considering each cycle in turn, Grigg could find no evidence of the operation of the positive check of higher mortality resulting from incomes falling below subsistence levels. Rising mortality may have ended the first two cycles, but this was probably exogenous caused by pandemic diseases, notably the Bubonic plague (see McNiel (1977), Hatcher (1977) and Hollingsworth (1973)). (The lack of influence of endogenous mortality is shown by there being little relationship between life expectancy at birth and real wages, even counter cyclical movements over the period 1551 to 1851 (see Schofield, 1983).) The virulence of the plague may have been in part density dependent, and hence a Malthusian response as abundance of people living in close proximity eased its transmission. Hatcher (1977) writes that the Bubonic plague in 1348 may have reduced the population of England by a third. Langer (1964) states that "in Western and Central Europe as a whole, the mortality was so great that it took nearly two centuries for the population level of 1348 to be regained". In the first cycle Grigg indicates that by 1300 the technological possibilities of medieval agriculture were largely exhausted and, as the population grew, the Malthusian negative feedback loop became apparent as incomes fell. However, the operation of the positive check in its most catastrophic forms was prevented by migration to underpopulated areas of the countryside and later by high sporadic exogenous mortality.

After this period there appears to have been a period of underpopulation as witnessed by land returning to long-term fallow and rising incomes. Thus, when exogenous mortality became less severe in 1450, the population grew and abandoned land was reclaimed (McNeil, 1977). By the 1620s in England, according to Grigg (1980), population pressure was manifested by high levels of rural poverty. The demographic response to this could have led to the establishment of the European marriage pattern of late age of marriage and a relatively high proportion never marrying - thus the use of the preventative check. Before the eighteenth century Goldstone (1986) argues that in England fertility fluctuations were a result of changes in the proportion ever marrying which changed in response to fluctuations in real wages with a twenty year lag. (Wrigley and Scofield (1981) suggest a longer lag.) Scofield (1983) comments, on seventeenth century England, that "intensity not only of marriage but also of sexual activity among the young in general responded to long-term trends in scarcity and plenty". Overseas migration and migration to urban centres also became important. Thus Hatcher (1977) considers there is a Malthusian cycle, but the cycle is interrupted by exogenous mortality.

After 1650 agriculture was improved and became more intensive with greater use of fertilisers, new machines such as the light swing plough, and the introduction of new crops, particularly turnips and clover.

Grigg (1980) writes that it is unlikely that these changes came as a result of population pressure as there was little increase in population between 1650 and 1750. Wrigley and Scofield (1981) point out that over this period mortality showed a slow secular trend upwards even although real wages were rising. This inhibited population growth, but they believe the chief mechanism was the preventative check which operated to give a "low pressure" solution to social and economic change.

In the 1740s growth began again and, as has been demonstrated in Section 2.7.1 of Chapter 2, for many populations the cycle is not yet complete. The main source of growth in England at this time (1751-1816), Goldstone points out, was increased fertility as he states that the gross reproduction rate rose by 31%, there only being a modest improvement in mortality. Most of Western Europe remained rural until 1850 yet the population grew very rapidly compared to past cycles. This was made possible by expanding the resource base in the agricultural revolution (particularly the introduction of the potato) and, as a result of land reforms, the drift of landless labourers townwards where there was employment in manufacturing. The emergence of proto-industrialisation disrupted traditional checks on fertility earlier and more universal marriage was possible.

This set in motion a "positive cycle", according to Wrigley and Scofield (1981), in which real incomes rose as fertility rose as a result of earlier marriage, and by children having attraction as investments as their costs were low and they could find work at an early age (Birdsall, 1983). Hence, the industrial age began.

However, there are cultural, social and economic differences between populations. In France, for example, fertility started gradually to decline before the end of the eighteenth century to a gross reproduction rate of 1.7 by 1850, a level not reached in England until 1900. In early eighteenth century France the positive check was, according to Wrigley (1969), the main mechanism keeping numbers and resources in balance. From this setting the preventative check became more important but not operating through delayed marriage but by fertility control within marriage. This may have resulted from the failure of the positive cycle to develop in France as the population stayed predominantly rural and land, since, unlike England there was no single-heir inheritance, was subdivided into smaller and smaller units and shared amongst a farmer's sons (see Birdsall, 1983). The positive loop also failed to develop in Ireland but so too did checks on fertility (although age at marriage did rise). The responses to lower living standards was to farm more intensively and make greater use of potatoes. The catastrophic appearance of the full consequences of the positive check was only prevented by large scale migration to the USA and to Great Britain, although there were around one million deaths attributable to the crises. (The French and Irish populations will be discussed again in Chapter 6.)

At this time emigration overseas was important in "drawing off" surplus European population (although it was of little importance in France).

In this latter cycle McKeown (1976) has shown that, with agricultural improvements and improved communications, diet improved (which may have raised fertility) and mortality rates fell significantly. This marked the onset of the demographic transition (Dyson and Murphy, 1985). Infant mortality was particularly improved which meant that parents could revise family targets downwards (Leibenstein, 1974). (Also of importance in lowering mortality may have been the gradual global warming which began around the end of the seventeenth century (Galloway, 1985, 1986).) McKeown believes that this mortality decline was the chief reason for the population rise in nineteenth century Europe. But this is not the only reason. Wrigley (1985) shows that in this period the life expectancy of the French population rose much more than neighbouring populations, yet the French population growth rates were much less.

The establishment in the second and third cycles of the Western European marriage pattern was aided by the household formation system (Hajnal, 1982, and Smith, 1981). In the late eighteenth and early nineteenth centuries in North-western Europe, young people left their parental home to become servants in another household where they served an "apprenticeship" in farming and, when they had saved sufficient resources, they married and bought a farm or took over their parents' farm (the parents opting for a retirement contract which guaranteed that the children would support them). Once the household was set up, the newly-weds employed servants and labourers to make their enterprise viable.

This system allowed the operation of a feedback loop with the economy, in that, in hard times, people were forced to stay in service longer and delay marriage in order to raise sufficient resources. Alternatively, if grain prices rose, the servants would receive more money and so allow earlier marriage. For this system to operate there had to be the possibility of late or avoidable marriage and independent non-extended households. These were present in England at this time.

Smith (1981) and Wrigley and Scofield (1981) point out that at least in England in the seventeenth and eighteenth centuries the system is more complex as there was considerable lag between changes in the economy (real wages) and household formation and hence fertility. The lag is of around 25-30 years before shifts towards a new style of behaviour became firmly established, and Smith (1981) suggests that this is due to the "experience of the older generation and the process of its communication to their children's marriage generation". Thus similar arguments are advanced as those Easterlin (1968) applies to recent populations.

Lesthaeghe (1980) adds to this discussion by arguing that the crucial mechanisms in North-west European demographic regimes that have allowed the establishment of the preventative check are the mechanisms such as household formation which influence the starting pattern of fertility. This is more robust than spacing or stopping as these depend on parental rather than communal control for enforcement.

Smith (1986) states that: "The rate of household formation was the process that underpinned or, in fact, determined the rate of population change" - this is suggested by the stability of household size and structure. Support for Smith's view is given by Wall (1979) by showing that in England there was low fertility between 1650 and 1749 when 18% of the population was in service, while between 1750 and 1821 there was higher fertility and a reduction of people in service to 10% of the population. This, Wall states, was not an age structure effect as the numbers in the age groups 15 to 24 years remained remarkably constant over the period 1541 to 1871.

To allow the operation of such a mechanism, Smith (1986) argues that the nuclear family must be supported through periods of "life cycle" crises such as on first setting up home, old age, illness and widowhood. This, Smith indicates, was achieved through monastic charity and parish support. Failure to provide such support would have led to the strengthening of the extended family and higher fertility, as is suggested by Cain (1982). The concept that, as "risk insurance" (welfare systems) increase, fertility falls as less family support is required, which is described by Cain, and other arguments based on wealth flows such as Caldwell (1982) suggests, are criticised by Smith (1986). Smith points out that:

(i) if marriage is late, then children may be too young to provide much support for parents;

(ii) if they marry young, then the children would be starting a family themselves and be too busy and have no spare money to give to parents;

(iii) there is no need to guard against widowhood in the European system as labour can be hired.

Cain (1983) does accept that, in North-west Europe where there is community support, it is "unlikely that the security asset value of children was an issue in the timing of fertility decline".

Thus "family organisation in pre-transitional England and North-west Europe relied on wealth flows in which those with few family dependants, who were economically active, gave to those with costly dependants or those who were economically inactive, regardless of their kin resources" (Smith, 1986). This system is very suscptible to age structure changes, and changes in real wages which encourage inter-generational competition for funds. Smith thus believes that inter-generational, intra-familial contracts are unlikely to be significant in lowering fertility.

Lesthaeghe (1980) suggests that community welfare systems make parents socially responsible not to have high fertility, encourage socially imposed moral regulation and acceptance that late age of marriage is socially responsible. Smith (1986) also refers to North-west Europe at this time as having a culturally determined "moral economy" as there was a socially defined norm of the basic minimum living standard, below which individuals are reluctant to descend when marrying and forming new households.

The emergence of proto-industrial activity giving opportunities for the young for a steady income away from agriculture disrupted the household formation system, and there was a move to earlier marriage, and hence higher fertility

(Goldstone, 1986). This was encouraged by a healthy economic growth and enclosures displacing people from the land. Thus, during the late eighteenth century, wage labour displaced servants.

Evidence of this process, outwith England, is given by Gutmann and Leboutte (1984) who show that in two districts of Eastern Belgium industrialisation was followed by a decline in the age at marriage and higher fertility. However, they also point out that the response to industrialisation depends on a variety of factors, the main ones being ownership of land, agricultural patterns and the length of time industrialisation takes.

From this brief review of agrarian Europe, it emerges that there is no straightforward density dependent mechanism acting on mortality and fertility. Although Mirowski (1976) did find some evidence of density dependence, a powerful influence on population change in early agrarian periods was exogenous mortality. (If foetal mortality is included in the positive check, then this check could be credited with some of the responses to price changes.)

In England at least there is evidence of a Malthusian system (chiefly the preventative check) operating prior to the nineteenth century. Scofield (1983) comments: "Before 1800, the situation developed much as Malthus had insisted it must: the faster the population grew, the faster the price of food rose, and the lower the standard of living fell, the grimmer the struggle to exist became. As Malthus had postulated, there were indeed long slow oscillations in the rate of population growth and in the standard of living." However, the mechanisms operated through social and cultural pathways as illustrated by the operation of the household formation system. If there was a relation of vital demographic rates to living standards, then these could well be overshadowed by other factors influencing prices at the time, such as overseas wars or climatic effects. Thus Smith (1986), referring to 1650 to 1815 England, states: "Rather than homeostasis the pattern suggests that fertility behaviour may have been responding systematically to secular change in its social, economic and biological milieu".

Density dependence might be more apparent in nineteenth century Holland, France and Ireland, but the social and cultural backgrounds make these populations exceptions to the general European pattern. More comment on France and Ireland is given in the next Chapter, where the findings of Chapter 2 are given support by showing that a sigmoid modelling framework is still appropriate for these populations.

In the current developing nations there is rapid population growth. Gwatkin and Brandel (1982) estimate that in the next century these populations will become stable. The main reasons for growth is falling mortality due to improved diet, medical and economic aid, agricultural improvements, particularly drainage schemes which help in combating malaria, technology transfer, pesticides and insecticides. (These are discussed further by Boserup (1981).)

Fertility also is high and is promoted by the early and nearly universal marriage rates and the presence of strong bonds between members in the extended family (McNicoll, 1984). (Thus industrialisation in Asia will not have the same effect, on increasing fertility, as in North-Western Europe (Goldstone, 1986).)

The existence of the extended family acts to prevent the

North-western European household formation system from operating. Hajnal (1982) suggests that in South-East Asia, as a consequence of the lack of such a household formation system, fertility is higher with households splitting when they become crowded. In such a system there is no buffer against resource shortage. Thus, if resources become scarce, the positive check emerges. Hajnal comments that, once the system is based upon joint households and the extended family, then it is hard for homeostatic mechanisms to emerge as stressful situations encourage higher fertility. This is also boosted by the principal mode of production being orientated around the family in which children contribute economically.

Lesthaeghe (1986) considers that there are signs of density dependence in sub-Saharan Africa where growth is setting new records and food production has taken a turn for the worse. Dyson and Murphy (1985) highlight the reduction of post partum sexual abstinence, and breast feeding as important in giving rise to shorter child spacing intervals and so higher fertility. (Reductions in sterility also contribute to higher fertility). In these populations, "the two positive checks, outmigration and mortality are slumbering just underneath the surface" (Lesthaeghe, 1986). To keep them there, Lesthaeghe (1986) and Grigg (1980) believe that there is need for increasing production and "grass roots" intervention by governments to promote fertility control. In some areas this has been very successful such as in the Chinese People's Republic (Yiu, 1980, and Chen, 1982) to give an extreme example. India has also had limited success with birth control programmes but on a less dramatic scale than the Chinese People's Republic (see Kirk, 1984).

Dyson and Murphy (1985) show that in Latin America and South-East Asia populations are undergoing the demographic transition to lower fertility, while African fertility is increasing, which may be the first stage of the demographic transition.

Caldwell (1982) reckons that the fertility fall will occur as a result of changing wealth flows which make the existence of the extended family less advantageous. Wealth flows within the family change as a result of a process Caldwell (1982) refers to as "westernisation" in which western values and attitudes emerge in these developing nations partly imposed and partly as a deliberate policy in the nations to hasten development. Crucial in this process is urbanisation and education. As these develop children participate less in family enterprises, become more expensive and begin to question economically supporting distant elders. Another important faction in westernisation is contact with developed nations, particularly through colonisation and technology transfers (see Boserup 1981, 1986).

Cain (1981) found that, for this process to work in societies based upon the joint household formation system, community risk insurance is required. If this is not present then high fertility will persist regardless of westernisation.

Cain (1983) points out that, in such societies, land and children can be considered as insurance. Cain (1986) states that cultivated land and fertility are positively correlated. However, children are the preferred form of security in environmentally hazardous regimes where land is less valuable and there is danger of losing the stock of resources. Also, if parents live longer than they anticipated, their stock of land could be insufficient, and weather is not the only threat to land resources as the land could be physically expropriated or lost as a result of fraudulent manipulation. Thus children are a safe complement to land resources and have the further advantages of easing work load and providing companionship.

Cain (1983) further argues that, in such systems, more than the minimum number of children are required for insurance as some die or default. The number of children desired depends, according to Cain, on:

(i) if male lineage is practised - more sons are required;

(ii) mortality rates or likelihood of serious disability;

(iii) risk of defaulting;

(iv) cultural restriction on women - if women are prevented from economic independence, they may have large families to ensure their own security;

(v) harshness of the environment.

Cain considers that fertility falls in the developing nations will not be smooth, as was the case in the West, but will be a two-stage process:

Stage I - mortality falls - so smaller family sizes are required to meet the desired target family size;

Stage II - the target falls - with environmental improvements or economic and social developments providing alternative sources of security.

Cain (1985) states that owned land can negate children as a source of security so there can be a negative effect on fertility. Alternatively, increased land ownership can create opportunities and demands for children, so increasing fertility. But Cain argues that land and children should not be seen as security complements rather than as substitutes. Robinson (1986) criticises Cain's risk hypothesis by pointing out that children are valuable to all families as a general asset, but are a poor investment for countering risk, and can find no empirical evidence in Bangladesh for the concept of "risk insurance". However, Cain (1986) argues that Robinson has misinterpreted his papers in that children as risk insurance is just one of a portfolio of determinants of fertility, and that Robinson has overestimated the range of options available to people. Attitudes to women are of importance and act to maintain high fertility in that male lineage and economic dependence of women on males still persist, encouraged in many areas by strong adherence to the Muslim religion. Daly (1985), taking a Marxist view, believes that, in Brazil, ruling elites act also to hinder fertility falls as they see themselves gaining economic advantage out of a large pool of cheap labour.

There are many other factors which alter population, mainly by increasing mortality, and are exogenous to the Malthusian system. For example, population problems (famine) in Ethiopia and Sudan, Lesthaeghe (1986) argues, result at least as much from high exchange rates, warfare, poor soil and severe drought (which might indicate a climatic change) than because the population is too large. Insect damage of crops and animals is also a major exogenous feature which keeps mortality high, especially in Africa. There are also problems for these less developed populations of destroying their own environments by allowing overgrazing and too much usage of trees for firewood, so contributing to desertification. Especially in Africa, exogenous mortality and disease act, and will act, as a major curb on population growth.

In the past, mortality crises were a major inhibitor of growth, and Hollingsworth (1973) describes India as suffering the positive check and high mortality resulting from acute famines, such as in Bengal (1769-70) and Shavi (1877-1878).

In China, the advance of civilisation has allowed there to be a substantial population since early times. This, according to Wrigley (1969), has been a "high pressure equilibria" where population development has been punctuated by long periods of famine. Over the period 1000 to 1700, the collapse of dynasties and high levels of warfare seem to have frequently checked and sometimes reversed population growth. (Famines were also frequent in China and India at this time.) This could have been the result of overpopulation and the operation of the positive check (see McEvedy and Jones (1978) who also think that Malthusian checks, predominantly the positive check (high levels of infanticide), operated in Japan over the period 1700 to 1825.)

Thus there is evidence of the positive check operating at early times. Lesthaeghe (1980) writes that "the central force in demographic homeostasis is the force of mortality". But there are clearly many factors exogenous to population size which affect mortality such as unfavourable environment, climate and outbreaks of epidemics. For example, in 1911-1920 the population of India fell as a result of twenty million deaths caused by an influenza pandemic (McEvedy and Jones, 1978), or the collapse of the native population of Mexico, which Hollingsworth (1973) suggests was due to social disruption and the introduction of new diseases as a result of the Spanish conquest of 1521. (An alternative explanation is presented in the next Chapter.)

On the fertility side there is no clear evidence of a density dependent reduction Social, cultural and religious factors seem to play a greater role.

Thus the review of agrarian populations indicates that, while density dependent mechanisms probably do play a part in population regulation, there are many other factors complicating the processes of change and regulation. These make a simple universal law of population growth that would underpin differential equations of sigmoids unlikely.

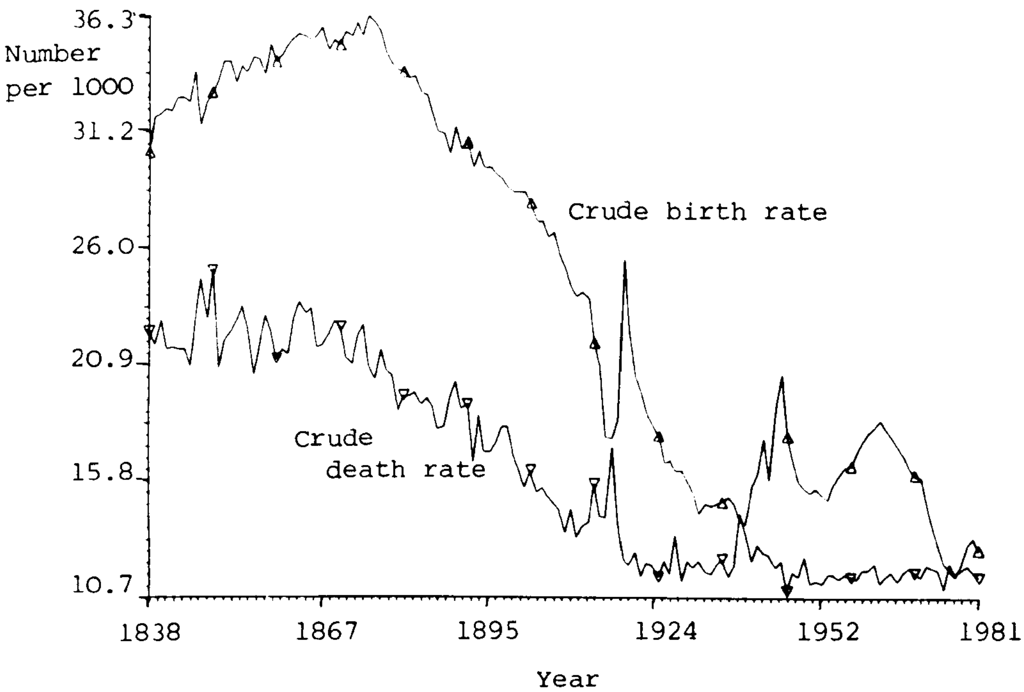
5.4.3 Industrial Populations

In the nineteenth century many Western countries became industrialised and the population changed from being predominantly rural to urban in nature. In this process a positive cycle developed in which both standard of living and population increased. Wrigley and Scofield (1981) found evidence of this cycle in late eighteenth century England. The progress of industrialisation was spurred by technological innovations. There is no evidence of the innovations resulting from population push. In recent times it would appear that technology and population are no longer linked in the way Boserup (1981) suggests is the case for agricultural populations.

In the early periods of industrialisation population growth came about as a result mainly of improvements in mortality although fertility also rose. Improvements in agriculture gave a better and more reliable food base and reduced mortality, although further urbanisation may have initially increased the number of deaths attributable to infectious diseases. However, the intervention of the state acted to combat such rises in mortality. For example, in the seventeenth century the imposition of quarantine achieved major successes in combating the bubonic plague. McKeown (1976) illustrates, changes in society made possible public investments to improve health, such as sanitation schemes, sewers, water treatment and an awareness of the need for public health programmes allowed the mortality fall to continue. It is only comparatively recently that direct medical intervention such as vaccination programmes, to alleviate such diseases as smallpox, has had any significant effect. (Vaccinations and medicines are more important in lowering mortality in the Third World and also contribute directly to increased fertility by reducing venereal disease-induced sterility.) In the West greater gains in increased life expectancy are probably more attributable to community health programmes such as anti-smoking campaigns.

At the onset of the industrial phase, initially fertility rose as the household formation system collapsed as increased urbanisation gave the young the possibility to leave the land for industrial wages, and hence economic independence at an earlier age. By 1871 in England fertility started to fall and so lagged the mortality fall (see Figure 5.4.3.1). The fertility fall can be explained in Caldwell's (1982) terms - when wealth flows between children and parents are no longer favourable to the parents. This occurs as children are less essential for risk insurance in an urban environment; space in houses would be at a premium so large families would have congestion problems; also increased education and acts of law (such as the 1819 Child Employment legislation in England) restricting child employment (see Birdsall, 1983). These and the non-familial orientation of production meant that children were no longer of such economic advantage as had been attributed to them under agrarian regimes

Figure 5.4.3.1: Crude birth and death rates for England and Wales 1838-1981



orientated towards the familial mode of production. Community risk insurance was also well developed so there was no need for children as protection against risk (Cain, 1983).

By the 1930s fertility and mortality had fallen to low levels and almost balanced, halting population growth (see Figure 5.4.3.1). Indeed, falling fertility gave rise to fears of population decline in Europe. Winter (1980) discusses these fears showing that they were focused on concerns of losing military might and racial supremacy, erosion of class differentials, and it was thought that high fertility was required to produce geniuses on chance grounds. However, after the Second World War there was a rise in fertility which, in some cases, gave rise to new cycles of growth (France and the USA are discussed in Chapter 6). In many of the Western populations examined in Chapter 2, the renewed growth is short-lived and fertility fell to sub-replacement levels. In the future fertility may oscillate around replacement levels in the final stage of the demographic transition (Bourgeous-Pichart, 1981).

Of the demographic components fertility is the most volatile and hardest to explain in the modern period. Various researchers, such as Wilkinson (1967) and Butz and Ward (1979), have allied fertility trends to cycles in the economy: when the economy is booming, fertility rises, and when the economy slumps, fertility falls. These theories are often associated with Kondratief waves, with periods of boom followed by slumps of approximately 50 years, or by some other periodic wave. These waves are discussed by Freeman (1983) and Rostow (1960, 1980). However, despite extensive statistical information and analysis, empirical validation has proved elusive, and one must note, as

Freeman does, that the theory of waves in the economy is inconclusive and of a speculative nature.

Fluctuations in post-transition fertility have been explained by Easterlin (1968, 1978, 1980) in a Malthusian way. Easterlin argues that periodic fluctuations in fertility levels are related not to long waves but to internal demographic factors of the population. For example, the fertility boom in the USA in the 1950s, Easterlin indicates, was a result of the relative economic advantage of young potential parents, due to their comparative scarcity in the population. Thus the young had more choice and power in the labour market and so could obtain at an early age good well-paid jobs and family building could begin early. Thus this generation has higher fertility, whose children, in turn, compared to their parents, find themselves economically disadvantaged as there is increased competition in the labour market. As a result, this group will act to delay marriage and restrict fertility (hence the preventative check).

Easterlin refers to the ratio between a couple's earning power and their material aspirations as their "relative income". Easterlin thinks that material aspirations are formed through experience and that this experience is the standard of living of one's parents. Thus Easterlin (1980) writes "The material aspirations of young adults are largely an unconscious product of the environment which is very largely shaped by the economic circumstances or income of one's parents". The young person's expectations of earning potential are, according to Easterlin, largely formed by early experience of labour markets and of working and are influenced by economic growth. Decisions about when to marry and family size depend on their judgments as to the possibility of fulfilling their material aspirations. Thus the young adult from a small generation will be able to meet his aspirations more easily than if he comes from a large generation, and his age at marriage will be younger and fertility higher.

This models a self-perpetuating process in which fluctuations (which can be large) occur around a stationary state (this is related to Lee's (1974) discussion outlined in Section 5.2). The Easterlin model provides an explanation of empirical and simulation evidence that suggests variability around the sigmoid path is greater when the upper asymptote (or statonary) is encountered. (If these fluctuations are great enough, perhaps a break from the old path to an escalated path is likely (see Lee, 1974).

Easterlin (1978) also relates social factors to his model such as achievement in school, mental illness, alienation and crime, arguing that negative aspects of these are associated with large cohorts.

Smith (1982) is critical of Easterlin's theory as:

(i) he could not find empirical evidence of the cohort specific fertility behaviour suggested by Easterlin. Smith considers that there was stronger period relationship;

(ii) he could not find evidence for Easterlin influences on the timing of births in relation to relative income. Smith considers the concept of relative income to be suspect as it is difficult for the young to equate their likely incomes with that of their parents when they were young;

(iii) the post-war American situation does not concord with Easterlin's hypothesis as, according to Smith: "It seems not be true, as Easterlin supposes, that the small birth cohorts entering the labour force in the 1950s fared especially well. Even despite concurrent increases in numbers and proportion of women seeking jobs, the larger cohorts entering the job market in the 1960s fared as well, perhaps better".

(iv) Easterlin's theory fails to consider the growing attachment of women to work.

Smith reckons that American fertility can be better explained without recourse to Easterlin theory. This model has not been very successful in explaining fertility changes outwith the USA. Many of the Easterlin model concepts could be culture specific so invalidating the general use of the model. Ermisch (1983) could find no evidence of an Easterlin-type process in Europe. Also the model does not reflect the important aspect of increasing economic independence of women, and the increased number of working wives, which might make young adults even from a small cohort less inclined to have large families. Ermisch and Joshi (1983) confirm empirically the strong negative effect on fertility of women's earnings, and that timing of births are related to having children when they are less costly. (This view is followed by Butz and Ward (1979) for U.S. fertility.)

Other flaws with the Easterlin approach are neglecting the influence of media advertising on aspiration formation, and technical and organisational changes in work disrupt the operation of labour markets. For example, skills gained over many years may become obsolete, so putting younger workers into a stronger position. Thus earnings are not necessarily dependent on cohort size. Also unemployment caused by structural changes in the economy or foreign competition might lower earning expectation independently of cohort size. Therefore, a closed demographic model as Easterlin's is inappropriate as fertility is influenced by many externalities.

Views possibly more suitable for contemporary developed societies are those of the New Home Economics model, such as advanced by Butz and Ward (1979, 1980) and Ermisch (1983), which are based on maximising a utility function whose main components are income, lifestyle and allocation of time within the family. These models are not density dependent as fertility changes in response to exogenous changes in the economy. The approach was originally postulated by Becker (1965, 1974), who particularly emphasised the cost of time in the labour intensive processes of child rearing, and is seen as having to be traded against other types of home production, including leisure, and so should be incorporated in income optimisation.

Ermisch examines two family types. One, where only the husband works, for which rises in economic benefit from work increase fertility rates and falls in real income result in restricted fertility in order that standard of living aspirations can be achieved. In the family where both spouses work, increases in income for wives make more women more committed to the workforce, with the consequence that fertility is restricted. Fertility is restricted by a short spell of closely spaced births, either early in marriage or delayed to later years. In the Butz and Ward (1979) static model, fertility moves countercyclical to economic changes as women choose to have children when they are less costly in terms of time foregone at work. In their dynamic model, Butz and Ward build in anticipation and show, with empirical evidence, that in the U.S. births are compressed into periods when the opportunity costs of children are expected to be low relative to later periods. They also found that increased prosperity caused potential parents to revise their expected completed fertility downwards, and targets can change substantially as new economic information is obtained. (In their model Butz and Ward suggest that, for men, the roles of parent and worker are compatible, but this is not so for women who must make "trade-off" between the two roles.) Ermisch argues that economic changes that raise women's wages may induce more women to work, but the effect is not symmetrical as, once in the labour force, then, in general, women will be reluctant to leave it, and, if wages fall, attachment to work may be stronger as wives try to preserve living standards.

To counter the suggestion that the rise in numbers of working women is the result of, rather than the cause of, fertility decline, Butz and Ward (1979) and Ermisch (1983) show that women have been "pulled" into employment rather than "pushed" as women's wages have risen relative to men's. However, Ermisch also recognises that there are some Easterlin-type effects in that, if a married woman works, then her offspring will have higher material aspirations and, more than likely, her daughters will try to remain in employment if they become married.

In the New Home Economics (NHE) model the affluence of the two-earner household allows the purchase of more leisure goods, such as foreign holidays, and more consumer products. These compete both in financial and on time allocation grounds with children. Thus children have associated with them a high opportunity cost. Therefore, fertility is negatively affected by attachment of women to the labour force and positively influenced by male employment. Changes in birth probabilities are primarily affected by period influences rather than by cohort influences as Easterlin (1978, 1980) would have it (Ermisch and Joshi, 1983). Therefore, if female work participation rates and/or their incomes relative to men's rise and male unemployment rises, then fertility will, according to the NHE model, decline. In the long run, Butz and Ward (1979) reckon that fertility will approach an asymptote from above. This will be a socially defined minimum level. (This concords well with the European situation over recent years, see Calot and Blayo, 1982.) However, Ermisch (1981) considers it likely, in the long run, that there will be a gentle secular rise in fertility fostered by one-earner families in response to economic growth, and so fertility will approach and asymptote from below.

There are exceptions to the model which makes its general use unsatisfactory unless cultural aspects are accounted for. For example, fertility in the Netherlands has fallen, similar to the rest of Western Europe, but female participation in the workforce is not very high.

As a consequence of greater work participation, women become more economically independent and, as divorce becomes more accepted, divorce rates have risen (possibly encouraged by break-ups over career moves to different geographical locations for the spouses). If there is heightened awareness of divorce in society, then perhaps there may be increased expectancy of divorce, which may encourage greater economic participation by the wife as a hedge against economic loss (this has been found by Greene and Quester, 1982), and consequently gives another negative factor on fertility.

With the growth consumerism (Bell, 1979) has created, alternative sources of pleasure (both service and manufactured output) in place of children, increased education into adulthood and the features of the NHE model has meant children have a high opportunity cost and so fertility falls. In the long run it seems likely that fertility will go to fairly constant levels, but at a socially defined minimum (which on present trends seems to be just below replacement levels) rather than a maximum (Woods, 1983). Butz and Ward (1979) suggest that in the U.S. major fertility rises can only come about if there is a large fall in young women's employment and a substantial increase in the supply of pre-school or day care facilities. The first is very unlikely as a result of structural changes in the economy expanding the service sector which has a high propensity to employ women.

The likelihood of obtaining such low fertility targets (rather than overshooting) is increased with improved contraceptive techniques and increased use of contraceptives. Fluctuations in births will reflect changes in age structure rather than attitudes or economic changes. Another feature of the process outlined above is a return to later ages at marriage after a phase of early marriage after the World Wars and in the early years of industrialisation, and thus a re-establishment of the Western European marriage pattern (Coleman, 1980, Kiernan and Eldridge, 1985). However, marriage in recent times has often been preceded by a period of "living together" (trial marriage perhaps) (Murphy, 1985).

Another factor contributing to lower fertility is decline of fundamental religious belief in Western industrial nations (Simons, 1977, 1986). Simons suggests that people's values are on a "relativist-absolutist" scale with strong absolutism corresponding with absolute, unquestioning acceptance of traditional religious doctrines, whereas relativism implies a greater questioning of traditional beliefs and a more hedonistic view (individual economic rationality). High fertility then correlates positively with absolutism and low fertility with relativism.

So, with declining religious values and more commitment to the economic system, societies' attitudes moved near the relativist end in North-Western Europe.

Perpendicular to this scale, Simons argues, is an individualism to collectivism scale. The argument is that the more individualist the society, then the less governments and pressure groups can do to change the values of society, and Simons suggests that, in Europe, there has been a move to more individualism.

By its socio-psychological nature this argument is fairly subjective. However, Simons (1986), using data from the International Values Study (1981) did carry out some statistical analysis and found that the countries with the highest fertility rates were to the collectivist-absolutist end. Unfortunately, Simons does not clarify the relationship between the various economic indicators and his variables by which he measured the scales. Thus, by employing cognitive dissonance theory, a move to relativism and individualism might be merely a reaction to economic changes, to bring attitudes in line with the action of reducing fertility, rather than vice versa.

Simons can also be criticised on statistical grounds as his perpendicular axes are likely to be strongly correlated. Lesthaeghe (1983) gives similar reasons for the fertility fall, arguing that both economic and ideology changes are important. Changes in ideology, Lesthaeghe considers, originated in the nineteenth century and have led to an individualistic orientation emphasising material doctrines. Lesthaeghe gives evidence of the attitudinal change as the legitimisation of voluntary childlessness, cohabitation outside marriage, non-conformist sexual behaviour, abortion, divorce, suicide, and small families being justified not by budget but by requirements to allow desired lifestyles.

Lesthaeghe (1983) writes: "The basic reasons for the association between economic growth and the change in demographic parameters of family formation, dissolution and reproduction are (i) that rapid increases in real income fuel individual aspirations, and (ii) that the opening up of new employment opportunities creates an impression of lowered economic vulnerability. This, in its turn, allows individuals to be more self-reliant and more independent in the pursuit of their goals, which ultimately stimulates self-orientation and greater aversion from long-term commitments".

Thus, it is not economic growth which reduces fertility but the resulting change of attitudes, and this is a lagged response as attitudes take time to change. The change has been to greater rationality in fertility decision making; this is discussed further by Beckman (1978).

Regarding development in South-East Asia, Oshima (1983) identifies the following processes as being important in leading to an economic and demographic transition:

(i) the emergence of full employment;

(ii) spread of mechanisation;

(iii) rapid rise in family incomes;

(iv) structural economic changes, movement to light engineering away from agriculture.

These have brought increased productivity and have contributed to greater income equality.

They have also resulted in associated changes in the need for increased education and large increases in the propensity of women to work, due to demand for a female workforce in new industries like textiles, which have raised the opportunity costs of children. Further impetus is given to lowering fertility in rural areas by the introduction of mechanisation, allowing small farms to be viable in the absence of child labour. Therefore, this illustrates once again that there is no straightforward universal explanation for population change.

This brief review of some of the explanations of demographic change in industrial populations adds further weight to the earlier findings that social, psychological and economic factors all interact, and it is unlikely that a simple relationship exists, such as, say, Notestein (1983) seems to advocate that in industrial countries, undergoing the demographic transition, raising living standards means reduced fertility. Relationships are more complex, and theoretical justification for using sigmoids cannot be found in identifying mechanisms that cause demographic change or act to preserve demographic homeostasis. There are too many mechanisms, many of which are subtle, culture specific, and are continuously evolving, for a simple general model to support the use of sigmoids.

Social and psychological factors appear to be very important, which makes simple density dependent relationships and arguments related to carrying capacity unsuitable as a base for interpretations of population change in industrial populations. Yet, overall, there are parts of Malthusian theory which are still relevant and it can be seen that the dynamics of economic change and social responses to change act in ways to promote and inhibit population growth and whose interaction still appears to constrain populations to growth along sigmoid paths.

This view is supported by Woods (1983) who argues that standard of living has continuously increased in Western Europe, which in agricultural times increased fertility by Malthusian feedback on marriage patterns. Then, with industrialisation, fertility is controlled within marriage in order to ensure continued rises in the standard of living (this, Woods argues, is governed by Marxist relations). Eventually, as in contemporary times, fertility falls to a minimum socially acceptable level around which oscillation in levels occur, and it is likely that neither Malthusian nor Marxist relations apply. However, the pattern of fertility fall which Woods describes does support the use of sigmoids.

To sum up, it is argued that Malthusian-type influences have operated and that there is differentiation between the modern West and the developing nations. This is evident in the greater prevalence of the operation of the preventative check in the West and more evidence of the operation of the positive check in the developing nations. This leads one to believe that sigmoid models are suitable for Western populations, and exponential models may be more appropriate for the developing populations. However, in all systems, exogenous factors play or have played a major role in influencing population development.

5.5 The Validity of Using Sigmoids

The proposition that sigmoids can be established as a law of population growth, cannot be supported. However, one can argue that the suitability of the sigmoids appears to have arisen as more than a fortunate combination of circumstances and that the use of sigmoid models in the manner used in this thesis is valid, since there have been periods (often long) of homeostasis and, in general, growth has risen as a result of some change disrupting the regulatory mechanisms. This change can be the consequence of internal demographic factors or factors unrelated to population size. When a new cycle of growth starts, the relative strength of inhibitors is decreased and promoters increased. In time new inhibitors eventually evolve, the power of growth promoters may weaken, and the population approaches stationary levels. It is when growth occurs that sigmoids are appropriate and it is suggested that cyclical models might be more appropriate when the population equilibria is approached.

Within each cycle of change the use of sigmoid models is adequate for both describing and forecasting population change. Adequacy cannot be judged on whether or not some theory supporting the model can be formed, but that empirical support is good and that sigmoids are as good as alternatives.

Since more accurate forecasts have been made than conventional forecasting (projection) procedures, component populations modelled and the methodology is simpler, less dependent on professional and end user judgement, there are advantages in using sigmoids. These models cannot be discarded on grounds of validity, for to do so would be to discard all total population modelling and forecasting procedures since, as this

Chapter indicates, none of the models developed so far can suit all populations at all times.

The lack of a simple relationship between mechanisms that alter population growth and sigmoids means that one cannot treat the demographic interpretation of the models' parameters with much respect.

It would have been useful if the upper asymptote of the sigmoid could be considered as likely. But, at this level, populations oscillate and one should not be surprised if, as stability is expected, populations grow beyond this asymptote and a new cycle might even begin. Further, it is not sensible to view the asymptote that arises out of the least squares fit of a sigmoid to a population as reflecting the population's carrying capacity, since most populations can fluctuate around the upper asymptote without encountering the positive check as this upper level is less than carrying capacity (although for developing nations it might be reasonable still to equate the upper asymptote with biological carrying capacity).

Thus one should not consider the interpretation of the models' parameters as evidence of validity (as Meade (1984) maintains, see Chapter 3). However, this is not to side with Feller (1940) who dismisses sigmoids as mere coincidental good fits, since the models can be interpreted in demographic terms (such as the relation to Malthus or Woods (1983)) and components of population can be inferred from the model of total population. Also the relation of births and deaths to a logistic process, as discussed in Chapter 4 does give mathematical interpretation of the demographic transition. Thus Lotka (1931) gave an explicit formulation of the transition process.

The discussion in this Chapter also advises caution for using sigmoids for long-term forecasting, although empirical evidence supports their use up to fifty years, provided there are no changes to a new cycle of growth. Thus there is agreement with Pearl (1925) who wrote of his logistic models: "The figures really express what the probable future populations would be if the conditions of the nineteenth century were to remain permanently unaltered. For the next ten or twenty years ..... the predictions given by the curves are undoubtedly highly accurate in most cases. But longer-range predictions could be taken seriously only by someone who denies the fact of any and all evolution."

Changes in the cycles of growth can be detected by considering how inhibiting and promoting mechanisms are changing. These have been manifested in fertility changes in recent periods (France, USA in the 1950s) and so can be detected when birth forecasts are compared to observations.

The problem of obtaining a general model for population and detecting escalations are the subject of the next Chapter.

Chapter 6 - Holistic Modelling of the Human Population System

6.1 Introduction

In the previous Chapter the main demographic theories were outlined and related to sigmoid models and it was noted that, as population evolved, so explanations for their development had to change. The aim of this Chapter is to consider, somewhat speculatively, the possibility of developing a general holistic framework that could be used regardless of the political, social, environmental, economic and demographic setting, to describe population growth, and to assess the suitability of using sigmoid models as models and to produce forecasts.

If one considers that the determinants of population growth, i.e. influences on fertility, mortality and migration, act in a dynamic system, composed of a milieu of complex interactions, referred to here as the population system, this system is made up of many subsystems whose boundaries are diffuse and often overlap. When attempting to describe such a system, immense difficulties arise to those using traditional methods such as block diagrams and differential equations. Due to the complex interactions and inter-linkages, formulation of anything more than a small subset of the system by practitioners of these traditional methods proves to be impossible.

It is a proposition of this thesis that modelling subsets of this system is not amenable to inference outside the strict tenet of assumptions that have to be made. Hence, extrapolations to other populations or to other time periods using the same framework will often not be possible.

Quantitative models of such a general nature must also be dubious as so many of the key variables are not easily defined and measured in quantitative terms. Traditional models such as block diagrams and feedback loops also fail when phase changes in population growth patterns arise as the abruptness of change often causes difficulty.

Thus it is the purpose this Chapter to consider the possibility of an alternative means of holistic modelling, that of Catastrophe Theory (which, despite its name, does not restrict itself to the dictionary definition of 'catastrophe'). To this end, the Chapter is organised as follows. Section 6.2 considers differential equation modelling to illustrate the inadequacies of traditional quantitative modelling. Then Section 6.3 outlines the methodology of Catastrophe Theory and summarises, as an example of the approach to be followed, the work of Renfrew (1978, 1979) dealing with modelling social change in archaeological societies. With this background a model is developed in Section 6.4 and applied to some populations discussed in this thesis in Section 6.5, and so there is overlap with the previous Chapter. This allows in Section 6.6 a reconsideration of Catastrophe Theory models as holistic models in demography. The Chapter concludes by reviewing briefly and speculatively some likely future scenarios for the population system.

By the ambitious nature of this Chapter, the discussion must remain speculative, but it is hoped that the presentation of the Catastrophe model will be useful to those who wish to fit quantitative empirical modelling into a wider framework so that its limitations can be better assessed and assumptions clarified.

6.2 The Differential Equation Approach

Systems of coupled differential equations have been successfully applied to model many complex, interacting processes in the physical sciences, ecology and in social sciences (see May, 1981, and Boyce and di Prima, 1977). The population system would seem a candidate for such modelling, particularly as subsets of the system have been modelled using differential equations.

For instance, Lee (1986) outlined a phase plane approach to model a subset of the population system at the macro level based on a synthesis of Malthusian and Boserupian theory to show the relationship between resources and population growth. The phase plane represents, graphically, the interaction of the differential equations, and Lee considers a two-variable case of a plane upon which coordinates are formed by population size and technology.

Lee represents an area he calls the "Malthus space" as a subset of the phase plane. In this the equilibrium population size depends not only on available resources but also on technology, the long-run equilibrium wage rate, the rate of surplus extraction and how that surplus is distributed. The principle of this model is that increasing technology allows Malthusian checks to be overcome and the population to grow, but, for technology to increase, the population must be of sufficient size to support the technology and not so large that there can be insufficient surplus to be generated for investment. In constraining the population below this maximum level, the preventative check is important.

To this Lee adds a "Boserup space" to depict the Boserupian notion that population growth fuels technological innovation, and the adoption of these innovations allows further population growth. Lee gives two possible formulations of this space which are:

(i) the case where a critical level must be reached before technological progress can occur, and

(ii) the case where technological advance occurs with even the smallest population.

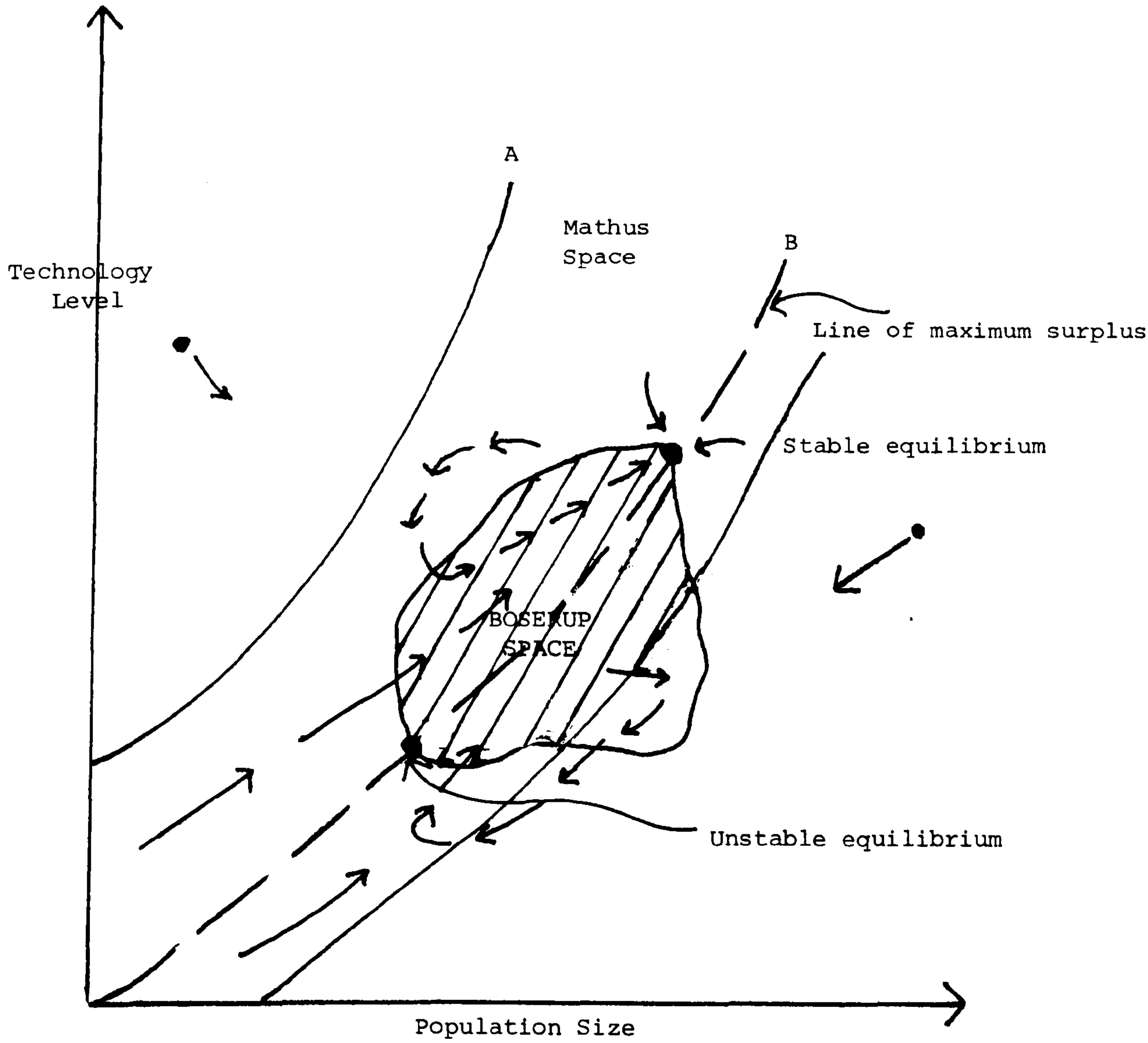
(Technological optimists, such as Simon (1981, 1983), would argue that the space is U-shaped suggesting that there can be unbounded progress and population growth.)

Lee forms a "phase plane" out of the interaction of the two spaces. Where the spaces intersect a node is formed. If a node is formed at the bottom of a Boserup space formed within a Malthus space, it will be unstable, whereas, if the node is at the upper end, the equilibrium will tend to be stable. Figure 6.2.1 illustrates this "dynamic synthesis".

As an upper node is approached, assuming income optimising the pace of both population and technological advancement slows and becomes stationary as real incomes fall. Lee recognises that such a system would be influenced by exogenous forces such as mortality due to viruses or plagues, foreign aid and climatic change. He considers it to be possible that there can be a succession of Boserup spaces within the Malthus space, representing technological and population growth within a hierarchy of major technological regimes.

In this model each regime is linked to the one above by a narrow "neck" or an isthmus, implying that the population would have to be suitably placed if progress was to occur and, to get to an upper regime, progress must be made through all other regimes.

Figure 6.2.1 - Malthus and Boserup Spaces Combined



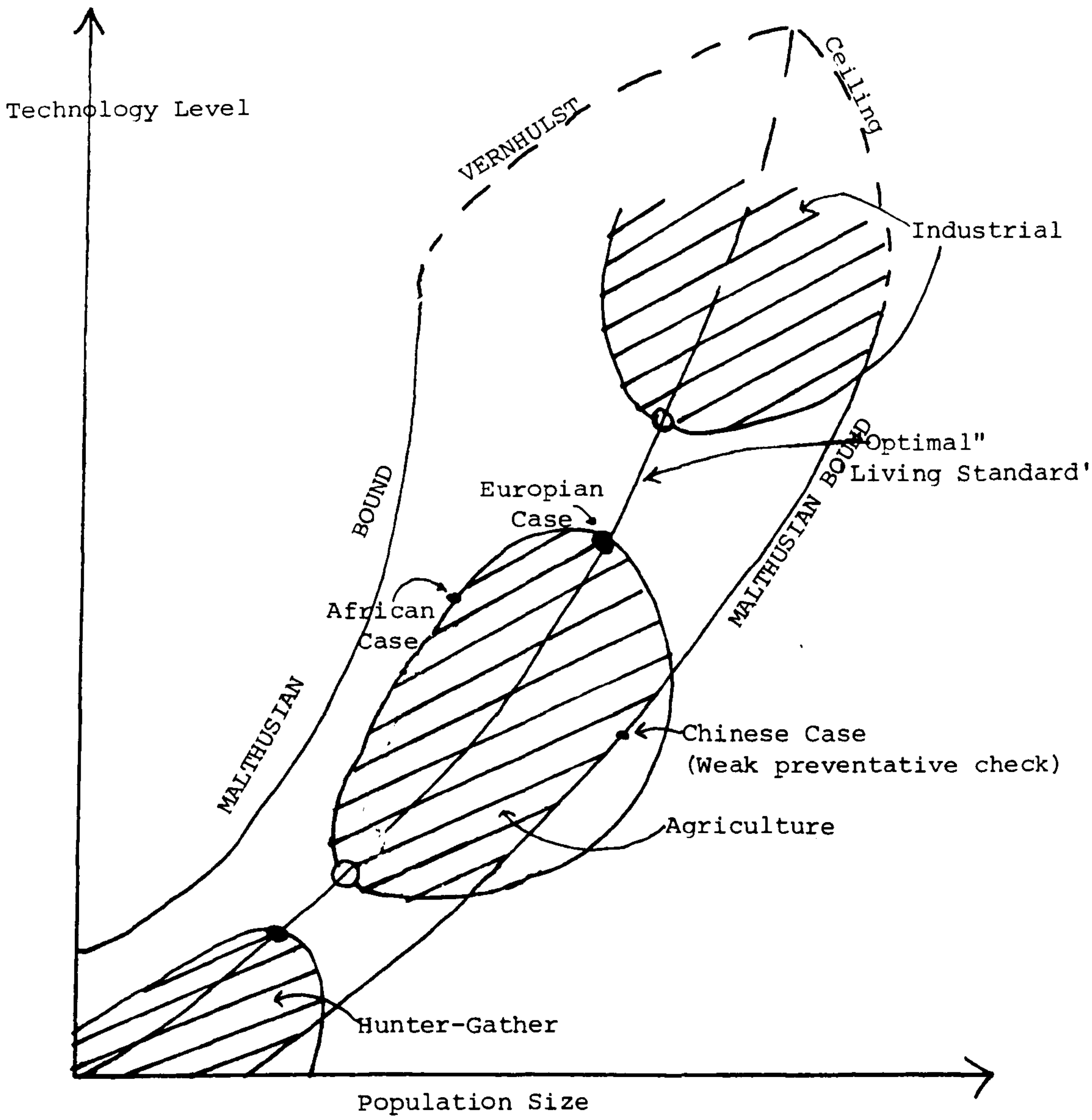
Hatched region is the feasible region for population

Arrows show the direction of population and Technology movement from different starting positions

(Alternatively, perhaps the regimes are unlinked, requiring discontinuous jumps in the population system to reach the next regime.) Such a construction is unbounded. Perhaps this should not be, and an upper boundary (a Verhulst ceiling) should be added to Lee's formulation, to illustrate the finiteness of the Earth's resources. This would result in the model illustrated in Figure 6.2.2.

Lee's model superficially appears successfully to encompass all possibilities and to help in explaining different population histories (Lee considers China, Africa and Europe). However, on closer examination the shapes of the "spaces" are rather arbitrary, the position of a population in these spaces can only be located rather vaguely, and many simplifying assumptions and value judgements are required such as how welfare or technology is to be measured. Lee's discussion of populations which conform with his phase plane model is in parts dubious. For instance, in Lee's framework, the poor situation of developing African societies is explained as being due to too sparse a population and too strong surplus extractive forces. This would imply that fertility control would change the state of the population of these societies to an even more primitive level, whereas mortality reduction would give rise to denser population, encouraging urbanisation, industry, infrastructural development, population and technological expansion and prosperity would increase. But many of the African problems are due to traditions of high fertility and mortality rates, poor shallow soils, native crops with low protein levels, the general inability of the land to support high densities, poor inexperienced leadership and tribal/racial divisions (Lesthaeghe, 1986, Caldwell, 1982).

Figure 6.2.2 - The Succession of Distinct Boserup Spaces and the Mathusian Space



O = Unstable Equilibria
● = Stable Equilibria

The hatched area is the feasible region for population

Hence it is difficult to conceive how increased population densities would improve the situation.

The relevance of incorporating Boserupian constructs is not applicable to present-day industrially developed nations where there no longer appears to be a direct interaction between technology and population. The nations with the fastest technological growth tend to be the nations with the slowest population growth. (This would suggest that Lee's phase planes are closed at the top end of technological development.)

Lee's framework also fails to give any aid in identifying mechanisms which operate in the population. The main criticism of Lee's model is that he uses a quantitative technique to give a qualitative description and so makes unnecessary demands on his methodology. The view followed in this thesis is that a quantitative model such as Lee's cannot be usefully developed for the population system.

In an effort for quantitative analysis, von Tunzelmann (1986) examined a subset of the population system by coupling population growth with "welfare" growth to describe movements of the English population between 1750 and 1850, using data from Wrigley and Scofield (1981).

Von Tunzelmann formulated the model:

$\dot{p} = -c + \alpha w$ ..... (1) $\quad$ p = population size

$\dot{w} = a - \beta p$ ...... (2) $\quad$ w = welfare

Coefficients c, a, α and β are non-negative.

This describes the "negative feedback loop" approximation to Malthus discussed in the previous Chapter. To incorporate a density dependence factor, von Tunzelmann adds $(-\gamma p^2)$, to (1) above, to give a system from which the logistic model results.

When this did not agree with observed data, von Tunzelmann focused on the Malthusian notion of moral restraint which suggests that there is a possibility that welfare can improve, yet population will not keep pace. Von Tunzelmann went on to test the models:

$$\left.\begin{aligned} \dot{p} &= a + bw - cw^2 + dw^3 - ep \\ \text{or } CNI &= a + bw - cw^2 + dw^3 + ep - fp^2 \end{aligned}\right\} \quad \text{and } w = a - \beta p$$

with all coefficients being non-negative.
The cubic term is necessary to give von Tunzelmann the required characteristics. The crude rate of natural increase (CNI) was used as an approximation to $\dot{p}$ and real wages as a proxy for welfare.

After fitting his model to population data, von Tunzelmann found a strong relationship for population change yet the phase plane he constructed was not supported, nor was the relationship between changes in real wages and population. To rectify this, von Tunzelmann suggested the incorporation of other variables such as poor law relief and corn prices.

There are some fundamental problems with such approaches of which the von Tunzelmann model is an illustrative example. For instance:

(i) Only a subset of the system is considered and hence many important influencing variables are not accounted for. In the above model, for example, factors such as technological changes, attitudinal change and physical wellbeing are not represented.

(ii) The models developed are used out of context. Von Tunzelmann, for instance, tests a Malthusian model for a period when there were changes away from an agricultural regime to an industrial regime with consequent changes in the regulatory mechanisms that he was trying to test.

(iii) There are data acquisition problems. For example, von Tunzelmann has to use real wages as a proxy for welfare, yet, at the time, there would have been much payment in kind (although it is widely recognised that real wages is the best information available).

(iv) The accuracy of the data in such a system is questionable and, in the above case, the data set constructed by Wrigley and Scofield (1981) has received criticism (e.g. Lindert, 1983).

To overcome these problems, the population system requires holistic modelling. This gives rise to immense problems of trying to:

(i) identify all the important factors which influence the population system;

(ii) identify how these factors interact;

(iii) construct the appropriate differential equations;

(iv) solve the system of equations to find the phase paths and equilibrium points;

(v) test the validity of such models.

To illustrate these problems, the construction of a set of equations that embraces many of the key factors is considered. These equations are presented only for illustrative purposes and it is recognised that there are deficiencies.

Rate of change of population size one can assume is related to the difference between current wellbeing (y_t) and wellbeing a generation ago (y_{tg}) when aspirations may have been formed. So it is relative wellbeing that is being considered, following an

Easterlin (1978) type of formulation. (This statement alone is open to criticism but will serve to illustrate the inherent problems of the approach followed.)

The rate of change of wellbeing may be formed out of three subfactors which are:

(i) Rate of change of physical wellbeing (y_p) which will include changes in life expectancy, literacy and infant mortality, and may depend upon population size, technological ability, affluence and environmental factors. The dependence on population may be one where y_p falls with increasing density of people, but rises as urbanisation becomes developed and falls again as cities become overcrowded. The simplest polynomial exhibiting these properties would be a cubic.

(ii) Rate of change of material wellbeing (y_m) - this may depend on population size in that too small a population precludes economies of scale and too large a population gives rise to resource allocation problems. The level of technology will also be important. (This also suggests a cubic form for the differential equation).

(iii) The rate of change of social-psychological wellbeing (y_{sp}) which measures the rate of change in attitudes, perception of the other elements of the system, group formation and individual needs. This may rise with increasing population but decrease as population pressure comes into play. Too sparse and too dense populations are thus considered undesirable (see Galle and Gove, 1978). Technological and environmental factors, such as the amount of carbon dioxide emissions (Broecker et al., 1985), may also be of importance. (This suggests a quadratic for the differential equation).

The other major factor may be the rate of change of technology (T) and this could depend upon the wellbeing of the population and population size. If the size of the population becomes too large, there could well be difficulties of surplus extraction and, if the population is too small, insufficient surplus could be formed and so hamper technological growth (again suggesting a cubic). Environmental factors also have influence.

From the above considerations, one formation of the key factors in the population system is as follows:

$$\dot{p}/p = a(y_t - \bar{y}_{tg})$$

$$\dot{y}_t/y_t \begin{cases} \dot{y}_p/y_p = -bp + cp^2 - dp^3 + eT + fy_m + \alpha E \\ \dot{y}_n/y_m = -gp + np^2 - dp^3 + jT & p = \text{population size} \\ \dot{y}_{sp}/y_{sp} = kp - lp^2 + mT + \beta E & y = \text{wellbeing} \end{cases}$$

$$\dot{T}/T = np - qp + ry_t + \gamma E \qquad T = \text{technology}$$

where constants a to r are ≥ 0; α, β, γ can be $\pm$ve.

This gives only one possible model of the system and gives only a restricted and crude abstract of the global population system, yet already the system outlined is too complex for useful analysis. Estimation of the coefficients of the above would be an unattractive proposition. This suggests that, for a holistic model of population, the differential equation approach is both inappropriate and is misleading. The reason for such conclusions are as follows:

1. The system is too complex to be able to include all possible interactions and one could never be certain that all factors and interactions had been identified. Also it is hard to assess the influence of factors which exist but have not been

included. Wunsch (1984) comments that "contrary to the physical sciences, for example, one is never sure in demography that one has included all the pertinent concepts".

2. Many of the variables are fuzzy and ill-defined such as social-psychological pressures, or the effect of environment and technology. Often quantification of these proves to be less than satisfactory.
3. The direction of causality changes as the system develops. (Wunsch (1984) refers to this as "fuzzy causality".)
4. It is hard to find measurable evidence for some of the constructs of the models so parts of such models are judgemental and arbitrary. Wunsch (1984) refers to the problem of multiple definitions and cites Kroeber and Kluckholm who give 52 different definitions of culture.
5. Testing the model is extremely difficult - how does one obtain the necessary data?
6. Sudden changes and discontinuities which lead to escalation of the logistic model are not easily modelled in this approach.
7. A succinct summary of the processes and assumptions is difficult.
8. In such a non-parsimonious system one must wonder how much the form of the equations is influenced by one's own ethnocentricities (see Ehrenberg, 1982).
9. The variables are made up of multiple factors which often act in contradictory ways; this makes model specification very difficult. Wunsch (1984) also points out that modelling difficulties arise as a consequence of people acting irrationally when faced with demographic choice, often being

influenced more by community norms than the desire to optimise individual utility, and so the sociological context of the model is important. (Also, in many cases, demographic events have a distribution of stimuli and responses and so require to be modelled statistically rather than the deterministic way this approach takes.)

10. The consequences, in terms of phase paths and equilibrium points of even slight errors in the above system can be dramatic. This is important as, at best, only rather crude estimates of the parameters of the model and initial conditions can be made. Thus there are possibilities of widely differing solutions.

Therefore, the conclusion is that, to model the intricacies of such a complex system, it is inappropriate to use differential equations as it would lead to too deterministic, inflexible, restricted and unparsimonious models of doubtful form to give a good base for modelling the population system. Wunsch (1984) gives further considerations that a demographic modeller has to face.

An alternative approach is to use the methodology of Catastrophe Theory (Zeeman, 1976) to give a qualitative description of the system.

6.3 <u>Catastrophe Theory</u>

Catastrophe theory is a mathematical method for modelling situations where there are discontinuous and divergent phenomena. This encompasses a modelling methodology of great generality and is considered by proponents of the theory to provide insights in situations where gradually changing forces lead to abrupt changes in behaviour.

It is basic to the theory that continuous changes in controlling variables cause discontinuous changes in the state of a system. Then it is possible to choose a coordinate system in such a manner that the types of discontinuities that occur can be represented by one of seven elementary catastrophes. These are listed in Table 6.3.1. In this theory the emphasis is on <u>qualitative</u> rather than <u>quantitative</u> descriptions of systems.

<u>Table 6.3.1</u>: <u>The elementary catastrophes</u>

Catastrophe	Control dimensions	Behaviour dimensions (response)
Fold	1	1
Cusp	2	1
Swallowtail	3	1
Butterfly	4	1
Hyperbolic	3	2
Elliptic	3	2
Parabolic	4	2

The most elementary catastrophe is the <u>fold</u>, which is shown in Figure 6.3.1. This Figure displays the equilibrium points of the state variable, X, and the control variable, "a". "A" is a stable equilibrium; if, from "A", variable "a" is increased, the system moves to "B" and, as "a" increases further, a discontinuity is encountered by the system and, if "a" is under 'pressure' to keep increasing, the only option is for the system to make a dramatic and sudden (catastrophic) jump to position "C". (A jump

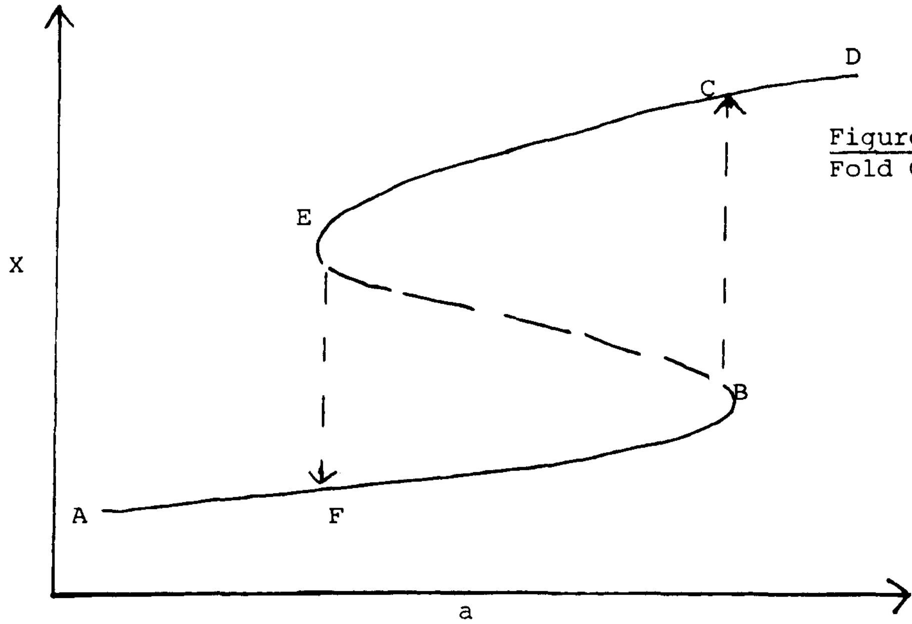


Figure 6.3.1 - The Fold Catastrophe

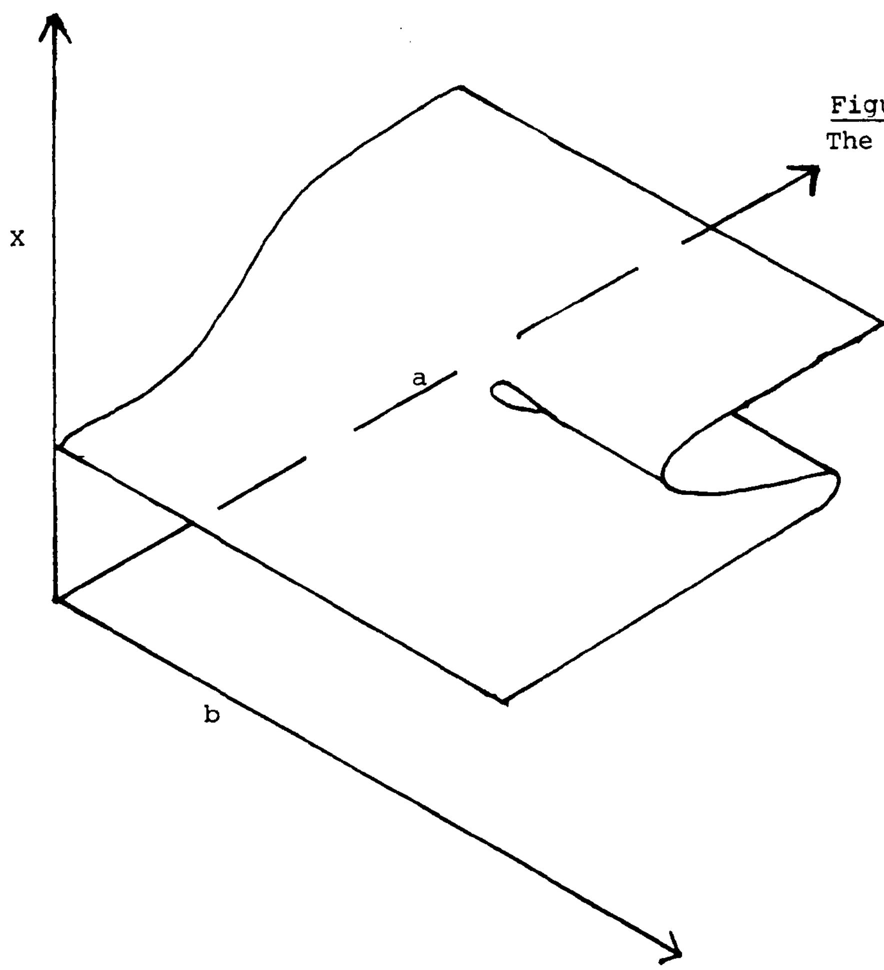


Figure 6.3.2
The Cusp Catastrophe

to a higher level is termed an anastrophe.) Further increases of "a" lead to a stable equilibrium, once again, when position "D" is reached. If "a" is now decreased, a different return path is observed - "CEFA". Following these paths two phenomena can be observed; these are:

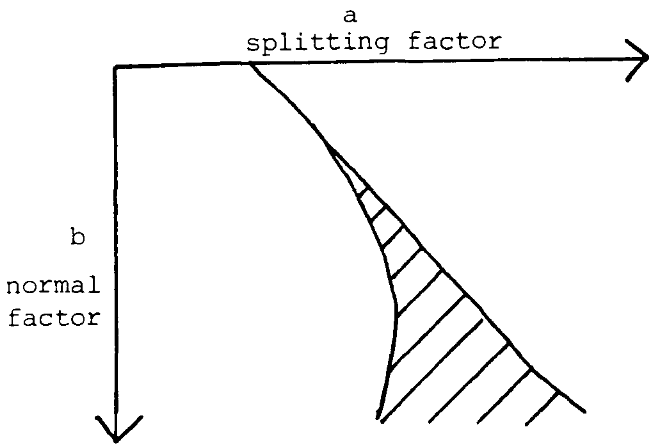
(a) bifurcation - a break point which gives rise to two divergent levels;

(b) hysteresis - the system delays the dramatic change in levels for as long as possible preferring to preserve the status quo. Eventually the pressure becomes too great and the catastrophe occurs, or the pressure diminishes and the system returns to stability.

The next level is the cusp catastrophe, the surface of which is displayed in Figure 6.3.2. This can be thought of as an extension of the fold to include another parameter "b". Until "b" passes a certain critical point, on the catastrophe surface, the system is fairly smooth and does not exhibit any catastrophes. It is when b and a change in a way to cause the response of the system to approach position "A" from below or position "C" from above that an anastrophe or a catastrophe will likely be experienced by the system. The cusp is shown as a continuous line. This is confusing as the underside of the pleat (B - E) represents an inaccessible area where the likelihood of observing the system in this state is close to zero.

This type of behaviour implies that the system has a control plane as in Figure 6.3.3.

Figure 6.3.3: The control plane for the cusp catastrophe



In this plane "a" is referred to as the splitting factor and "b" as the normal factor. The shaded area of the pleat of the cusp, which is characterised by two possible modes of behaviour, contains the bifurcation set. (For the simple fold catastrophe, the bifurcation set is marked as a point on the control axis.)

As more control dimensions are incorporated into the model the model's surface contains more discontinuities and has a more complex surface where there can be multimodal states for the system.

Modelling, using catastrophe theory, has been found to be useful in both the physical and social sciences. Poston and Stewart (1978) give many examples of catastrophe theory in physics, where some degree of quantification is possible. Zeeman (1976, 1977) applies catastrophe theory to a wide range of systems and examines many cases in biology, ecology and the social sciences where he suggests that there is greatest scope for the useful application of catastrophe theory and could provide a mathematical qualitative language for the hitherto "inexact" sciences (Zeeman, 1976). Zeeman (1977) attempted a quantitative catastrophe theory based description of a prison riot and of industrial disputes, but the quantitative measures of the system are rather dubious and Zeeman (1977) has received criticism (Zahler and Sussmann, 1977) for the arbitrary modelling and misleading conclusions. Indeed Wright (1982) suggests that, where quantifying is possible, then it is better to use alternative analysis. Zahler and Sussmann (1977) would agree with this when they criticise catastrophe theory on the grounds of confusion about continuity, inability to justify extrapolation, prediction being contrary to fact, lack of testable predictions and there being better alternatives. Zahler and Sussmann also criticise practitioners of catastrophe theory on their misuse of generality and mathematics, careless discussion of evidence, unreasonable or ambiguous hypothesis and spurious quantification. Zahler and Sussmann are criticising catastrophe theory as a quantitative model which would be used to produce quantitative predictions; in this sense their criticisms are justified. However, the view taken in this thesis is intended to build on that of Thom (1983) that catastrophe theory should be used as a language with which to explain actions. Hence, Thom states that for modelling "soft systems" catastrophe theory is not meant to be rigorous nor conceptually nor numerically exact. Hence, the use of catastrophe theory in this thesis should be judged as a qualitative model, and consequently Zahler and Sussmann's criticisms do not apply. In this light, Thom shows that quantitative prediction is not possible unless used in conjunction with a supplementary hypothesis. Even if Zahler and Sussmann's criticisms are justified, Guckenheimer (1978) argues that catastrophe theory models should not be disregarded as "they aid to organise our thinking and explore the implications of our assumptions".

The approach followed in this thesis then is that a catastrophe theory model will be constructed to give a framework for obtaining qualitative predictions that could be used to assess the suitability of the supplementary hypothesis. The supplementary hypothesis is that sigmoid models adequately describe population movements. If the supplementary hypothesis is found to be suitable, then quantitative prediction is justified.

A catastrophe theory model of this type of the population system would allow many of the factors which influence population growth to be "massed" into a few controlling dimensions. The control dimensions could be thought of as vectors of the influencing factors. Thus a parsimonious and a holistic description of the population system could be formed, and assumptions could be stated with reference to a catastrophe theory model (or catastrophe map).

To illustrate this description of catastrophe theory, and also to describe the methodology which will be used when populations are modelled, a catastrophe model of the collapse of early civilisations which was formulated by Renfrew (1978, 1979) will be described.

The archaeological record of several ancient civilisations, such as those of Mycena around 1100 BC, the Indus Valley around 1800 BC, and the Hittites around 1200 BC, have exhibited sudden dramatic collapses. Renfrew considers the case of the Maya civilisation's collapse which he described as leading to the onset of a dark age.

The symptoms of the collapse Renfrew lists as:

(i) failure of elite-class culture - manifested by the abandonment of administrative and residential structures, such as palaces, failure to maintain funerary monuments, temples, records and the abandonment of rituals and manufacture of luxury items;

(ii) the apparent rapid depopulation of the countryside and ceremonial centres;

(iii) the relatively short period of occurrence - from 50 to 100 years.

Explanations for such collapses have been in terms of natural disasters such as tidal waves, volcanic eruptions or earthquakes, or plagues or invasion or internal revolt. Such explanations may be useful for some civilisations (for example, Pompeii). Renfrew is sceptical about the general application of such explanations.

Instead, Renfrew advances a cusp catastrophe model with Degree of Centrality (C) as the behaviour parameter. High values of C are associated with strong kingship, the maintenance of a bureaucracy and large investments in palaces and temples. The control variables used by Renfrew are Investment in Charismatic Authority (I) and Net Rural Marginality (M). In Renfrew's model, I is a measure of the energy assigned to cultural devices to promote adherence to central authority. This involves three factors: direct and implied coercive force, dispensation of material goods and dispensation of symbolic benefits. Net Rural Marginality refers to the economic balance for the rural population. This consists of positive elements, which are the economic output of the population, and negative factors, which are the costs individuals pay to belong to the state. When the population is in a difficult marginal environment, M will be high even if the tax burden is low.

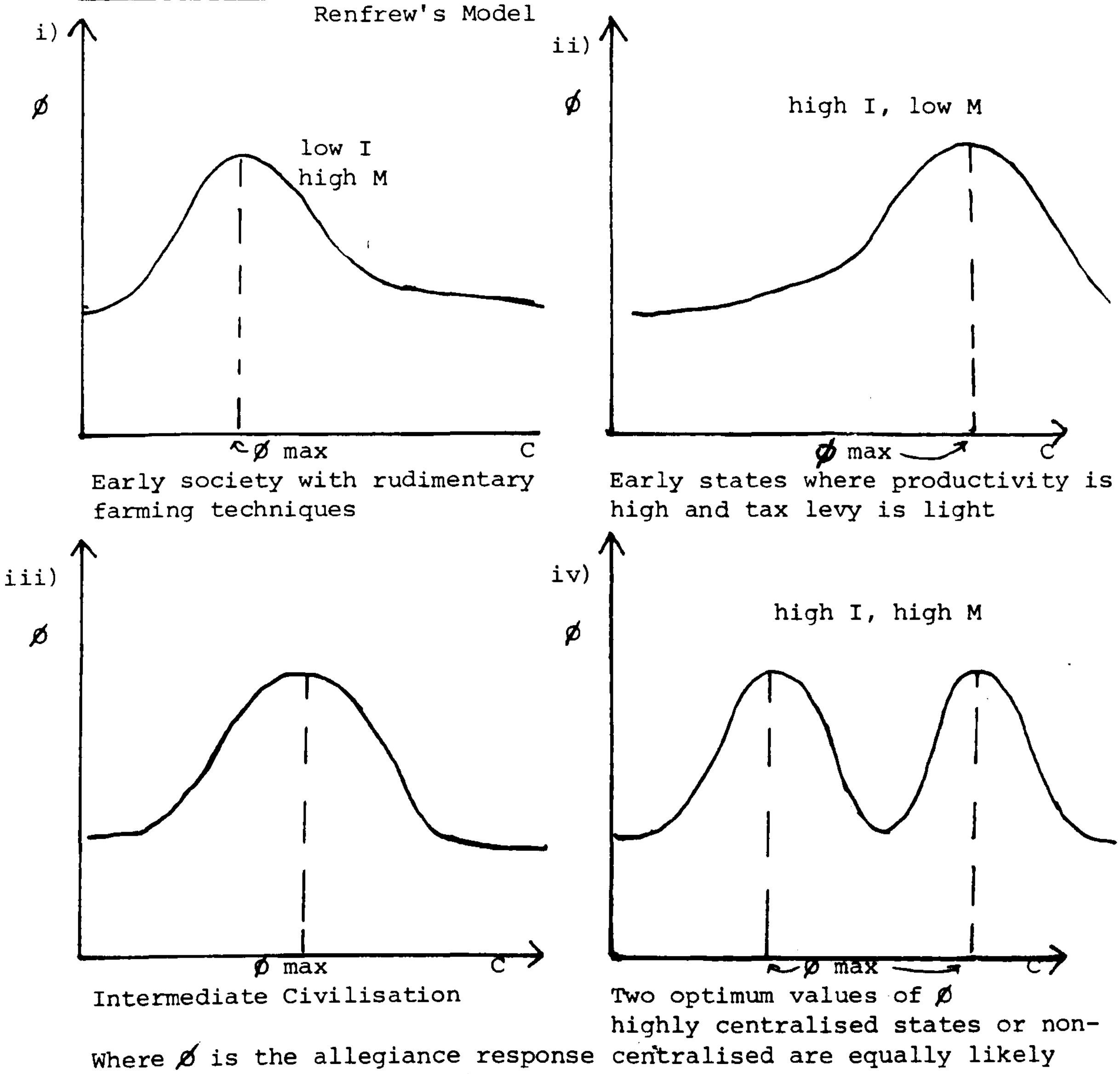
From this Renfrew suggests that a function (f) of C, I and M measures the wellbeing of the rural population. In this model the rural population acts in a way to maximise f locally, and the degree of Centralisation is seen as depicting the success of the central authority in commanding or attracting the adherence of the rural population to the central system.

Renfrew constructs the model by considering extreme cases of the control variables. The construct is illustrated in Figure 6.3.4a where ϕ is the allegiance response. From this Renfrew detects the presence of at least one cusp and this is illustrated in Figure 6.3.4b.

Renfrew gives what he describes as a typical system collapse trajectory for Maya or Mycenaean civilisation. This is sketched in Figure 6.3.4b.

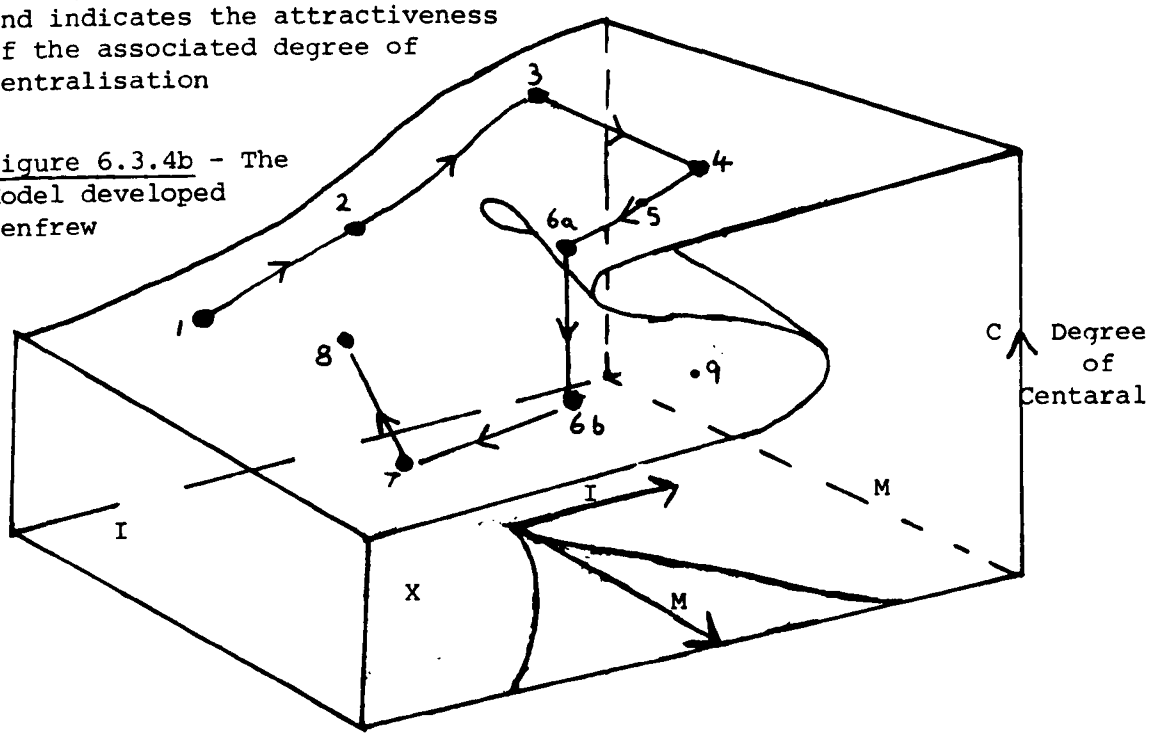
The degree of Centrality increases smoothly (1 - 3) with increasing I, until M rises and the tax burden increases on the population. As M rises further, central authority becomes less desirable (4 - 5) and I declines. When point 5 is reached, in terms of optimisation, point 9 is also a maximum, but represents a much lower degree of Centrality. As faith in the charismatic authority (I) diminishes further, and the local maxima of (5 - 6a) completely disappears, the system changes rapidly to 6b - 7, the central authority collapses and the population becomes more dissipated, small clans, tribes or family groups become the dominant type of social organisation. Renfrew (1979) extends this model into a Butterfly catastrophe to model the distinction between tribe, chiefdom and state. Renfrew (1978, 1979) also

Figure 6.3.4a - Extreme Cases of the Control Variables used in Renfrew's Model



Where Ø is the allegiance response and indicates the attractiveness of the associated degree of centralisation

Figure 6.3.4b - The Model developed Renfrew



develops a catastrophe model to show how slowly changing variables representing the mode of cultivation and rewards from communal life alter the attractiveness and suitability of settlement size and spacing, and can result in rapid and dramatic changes from nucleated to dispersed settlements (and vice versa). (It should be noted that there are alternative views of the Mayan population collapse, such as Zambardino (1980) or Hollingsworth (1973).)

Renfrew (1978) considers catastrophe theory to be useful methodology for modelling social systems, stating that "Catastrophe Theory has been applied in the past few years to a number of different cases. In each it has increased the facility with which the problem can be grasped. At the very least it offers a handle, a way of visualizing the rather complex interactions of multivariate phenomena".

From this discussion it is apparent that catastrophe theory offers a methodology to model complex systems where quantification is not appropriate. Zeeman (1977) illustrates the strengths of this methodology by stating: "Thinking by numbers using statistics and computers is analysis, while thinking by pictures, drawing the graph and identifying the catastrophe is synthesis. Synthesis is more important because it gives us concepts which we can grasp (or see) and upon which we can build further".

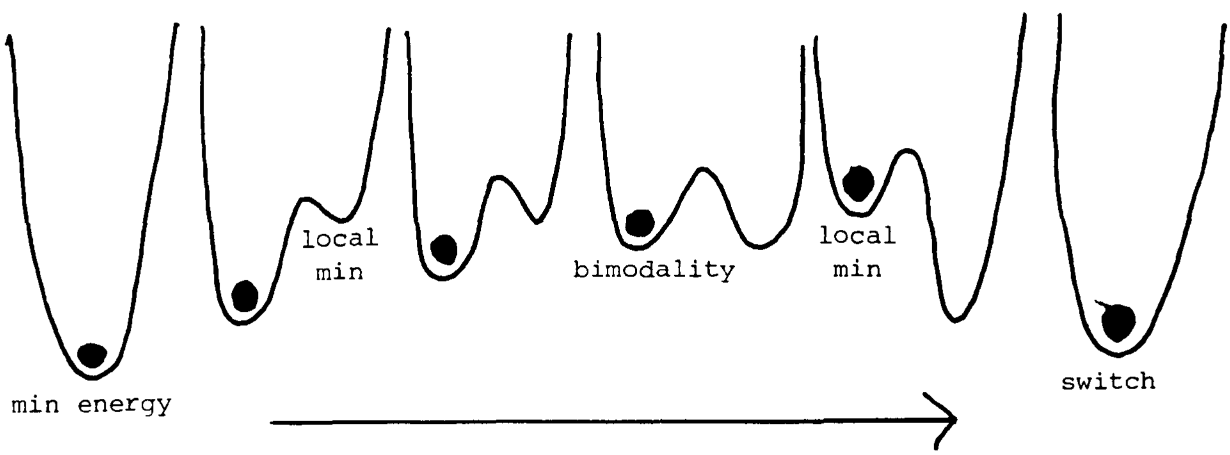
For catastrophe theory to be suitable for modelling the population system, the system must satisfy the required conditions of at least two of the following being present in the system:

(i) Bimodality - there are regions of the system where the most likely outcomes are either rapid population increase or rapid population decrease (or remain stationary). An example of bimodality from demographic literature is given by Ermisch's

(1979) review of Butz and Ward's (1976) New Home Economics model. He found that there are two classes of family, one where the wife has a low participation in the workforce, and the other where wives have a high participation in the workforce. Very few have participation rates near the median.

(ii) Hysteresis - this is explained in general terms as the system delaying for as long as possible in a position of a local minimum after bimodality occurs before switching to a lower level. This is represented pictorially in Figure 6.3.5 in terms of potentials for the model.

Figure 6.3.5: Model energy potentials



Examples of this type of behaviour in the population system are the perseverance of traditional lifestyles and desire to maintain the status quo. A more pertinent example, relating to the discussion in the previous Chapter, comes from the New Home Economics (NHE) model. If the economic advantages that women accrue from work external to the home decline, and so the cost of children is reduced, the working wife would, according to the NHE, delay leaving the workforce in order to try to maintain economic status. If the economic decline continues, reversion to economically dependent status could well occur quickly with consequent rises in fertility.

Smith (1981) also draws attention to inherent delays in the population system by reporting of a fifty-year lag between economic movements and changes in fertility.

(iii) Inaccessibility - this refers to areas of the behaviour surface which are inaccessible (or very unlikely, if the surface is thought of as a probability density function) such as the underside of a fold. An example might be a nation with a stationary population where its government institutes a fervent anti-natalist policy. The most unlikely outcome would be for fertility to remain unchanged, the likely outcomes being in line with government wishes and a fertility fall or an increase in fertility as a "backlash" to the government's policies. Thus it is an inaccessible region fertility to remain unaltered.

(iv) Sudden Changes - sudden, relative to past movements in the population system have been empirically noted in Chapters 1 and 2, in that there can be dramatic changes in fertility (e.g. France in 1945-1950, USA in the 1950s) or in migration and mortality patterns (e.g. Ireland in the mid 1840s).

(v) Divergence - small changes in the control dimensions can cause large divergent changes in behaviour. Examples of this must wait until control dimensions have been identified, but one may postulate as to why, for example, the French escalation of population growth occurs after the Second World War, yet ~~demographic-economic situation was similar~~ after the First World War ~~where~~ no escalation was found to have occurred.

Therefore, the population system does seem to have qualities which would allow the use of catastrophe theory as a modelling framework. Wright (1983) emphasises that there must be equilibrium states. There clearly have been, as illustrated by Chapter 5 and world population history.

Wright also emphasises that there must be a considerable difference in time constants between the behaviour and the control dimensions. There do seem to be appreciable time differences as behaviour can change quickly as illustrated by the doubling of fertility in France in the period 1944 to 1950, while change in possible control dimensions has been slower. The economic system, for instance, may change over cycles of fifty years (as suggested by Schumpter, 1964), and major cultural changes take some fifty years (Allen, 1985) or longer (Leach and Wagstaff, 1986).

6.4 The Control Factors and the type of Surface

Having shown that a catastrophe model may be both useful and appropriate to describe movements in the population system, the type of catastrophe model which gives the most suitable description will be considered in this section. Suitability is judged by the criteria of comprehensive description and parsimony of control dimensions. Before constructing a model, it should be pointed out that the model is general and does not just refer to discontinuities.

The first element of the model to be identified is the response dimension that forms the behaviour axis. Since it is population change that is being modelled, this variable could be the rate of change per year or total fertility rate (plus migration). However, although successful in physical systems, frequently in "soft systems" to take such a dimension of this type would be out of context since following the methodology of catastrophe theory restricts one to construct the behaviour dimension in the form of different potentials for outcomes. Thus the behaviour variable could more appropriately be thought of as the <u>mode</u> of a distribution of possible outcomes and reflects a probability distribution of the likelihood of a particular population to change in a particular direction. To avoid confusion with existing statistical terminology, this dimension will be labelled the propensity for population growth (this is not to be associated with Malthusian terminology).

To an extent the usefulness of catastrophe theory, as with any forecasting method, depends on the spread of the probability distribution around a central tendency which, in this case, is represented by the surface of the catastrophe map. This is is difficult, if not impossible, to measure. However, one can make the general statement that one will be more confident about the outcomes towards the extremes of the control plane away from the bifurcation point. Therefore, the spreads around the surface will be narrow at these points, compared with areas close to the catastrophe of the control plane where there will be complex interactions. These complex interactions result in a greater range of possible outcomes and more uncertainty. Figure 6.4.1 illustrates these points.

The construction of the control plane is now considered starting with the simplest case of only one control factor. If the simple "bottle" explanation of population change is used, then the control factor could be carrying capacity. This would give a description of the population system which would depict dramatic changes in behaviour as a consequence of a slowly changing control factor, and so would give the fold catastrophe model (this system is displayed in Figure 6.4.2).

Changes in carrying capacity can be simulated mathematically and confirm the possibility of catastrophic changes (see Lee, 1974, as discussed in the previous Chapter). It is conceivable that, when the "bottle" expands (i.e. an increase in carrying capacity), the population takes some time to respond. Perhaps people want to be sure that the improvement is permanent and to assess the magnitude of change and gradually to alter cultural orientations that suit the lower level of carrying capacity. If the improvement continues, then there may be a considerable increase in unused resources and the constraint of carrying capacity "retreats into the distance". The population may react in a manner similar to the colonisation of a new land and there

Figure 6.4.1 - Distributions around the Catastrophe Surface

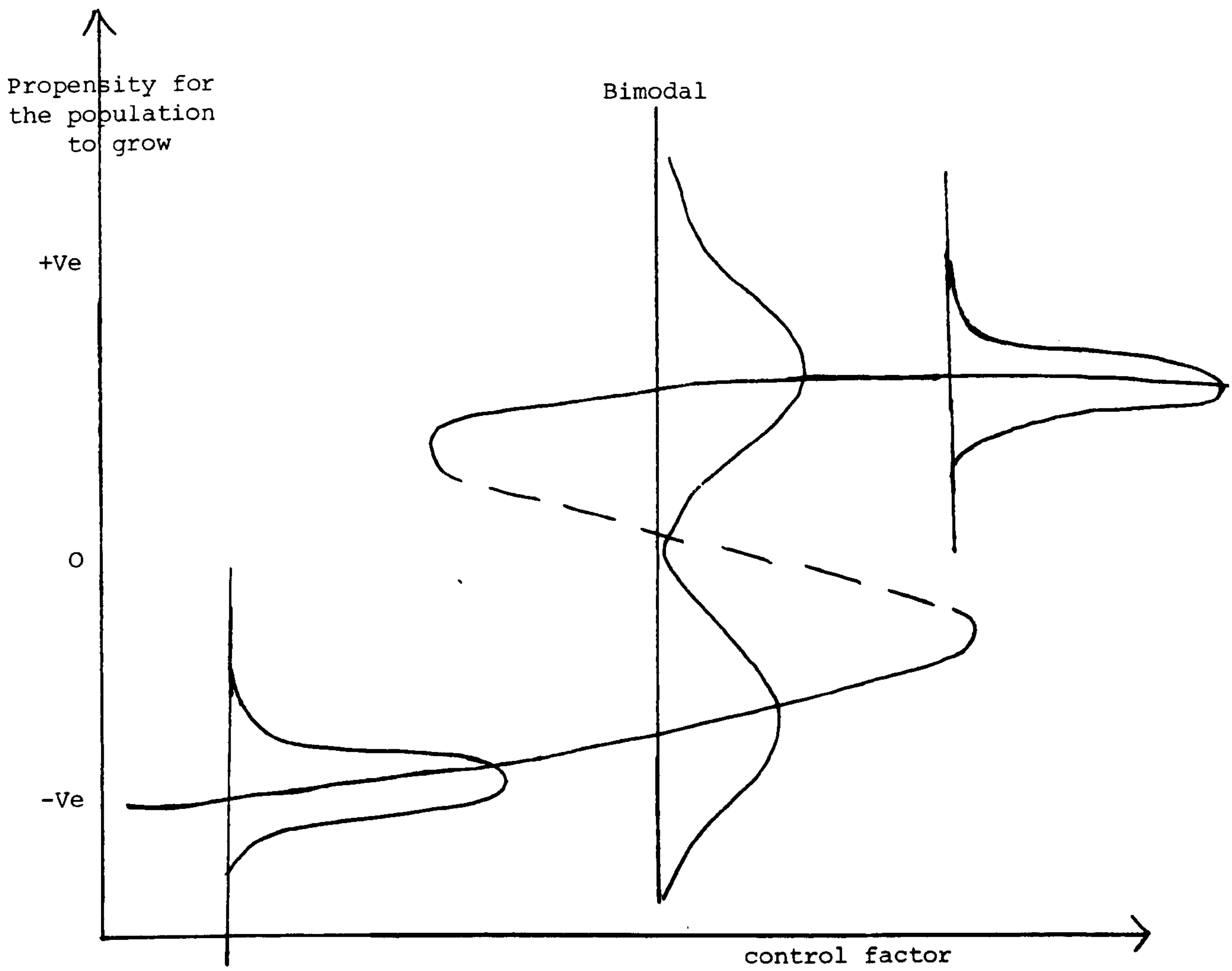
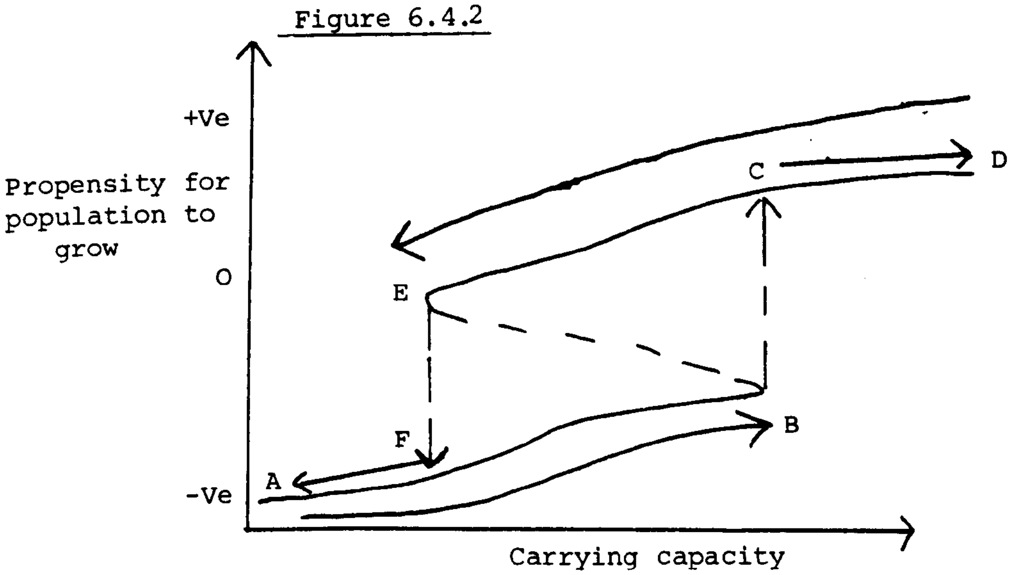


Figure 6.4.2



may be a considerable increase in the total fertility rate (combined with reduction in mortality rate).

Alternatively, if the population is growing steadily or is stationary and the level of carrying capacity falls, the desire to maintain the status quo, perhaps viewing children as protection against falling living standards and insurance against the impending "harder times", will maintain the existing growth pattern (or even lead to a short-term increase). If carrying capacity continues to fall (encouraged possibly by the failure of the population to employ corrective measures), the means of subsistence will no longer be sufficient and the Malthusian positive check will cause the collapse of the population.

If one was using such a model in conjunction with forecasts, and observed a population at position A in Figure 6.4.2, one would place a high degree of confidence in a forecast which predicted a move to zero growth with improving carrying capacity. But, once the fold is approached, confidence in such a forecast would diminish, and one would have to recognise the possibility that the population could remain unaltered in its propensity to grow or an anastrophic jump to a higher growth rate could suddenly be made. So it would be sensible to greatly shorten lead times for forecasts until the system moves away from the area of the fold. This model suggests that the position of the fold can be detected by monitoring the growth rate of the population, suggesting that bifurcation is likely when the growth rate tends to zero.

From Figure 6.4.2, it is also apparent that the path around the system is A to B to C to D, when carrying capacity increases, which is different from the path D to E to F to A, which occurs when the level of carrying capacity decreases. The anastrophe that occurs with the path A → B → C may be useful as representing population growth which may have occurred as a result of changes in social organisation and technology and resource diversification amongst Stone Age man (see Hayden, 1981). Another example is the growth of the Dutch population which occurred partly as a result of land reclamation. The catastrophic path D → E → F, on the other hand, may explain the Irish situation of 1841 to 1851 when the population declined from just over eight million to just over six and a half million. One explanation is that successive poor and failed harvests, likely due in part at least to population pressure, reduced the carrying capacity to such meagre levels that the population was forced into decline.

In this model the lower part of the surface can be thought of as representing a population where the various checks on population growth operate and are effective, whereas the upper part is where these restricting checks are either overcome or are relaxed. However, as has been discussed in the previous Chapter, such a simple description, based upon the notion of an expanding bottle, fails to explain the subtle and complex mechanisms that act upon the population system and why populations do not respond directly, and in the same direction, as changes in carrying capacity. Therefore, one must consider more complex surfaces.

The Cusp catastrophe model which is described by two control dimensions is now investigated. If the population system can be modelled with two control factors, then it is attractive to postulate that the factors operate at different levels, that of the society as a whole, and at the family level. The following list illustrates some of the main elements of the population system that influence population growth. These have been grouped into two levels:

Society level	Family level
Standard of living (Woods, 1983) Nature of technology (Boserup, 1981) Means by which the society is organised (e.g. feudalist, capitalist or communist) (Smith, 1986) Influence of religion (Simons, 1986) Physical quality of life (Hicks and Streeten, 1979) Wealth of country in both land and manufacturing, and degree of development in world ranking terms (Notestein, 1983) Amount of provision the society makes for its membership in times of hardship Environmental factors such as pollution (Broecker et al, 1985) and possibly climatic change (Lamb, 1982, Wrigley and Scofield, 1981)	Power that the family has Family type (nuclear or extended) Household formation (Hajnal, 1982) Need for risk insurance (Cain, 1983) Perception of the economic value of children (Caldwell, 1982) Wealth flows within the family (Caldwell, 1982) Psychological factors relating toneed satisfaction (Maslow, 1954) and to perception of the population system and the interpretation of societal norms Wealth flows between families (Smith, 1986) Loyalty to lineages or class (Renfrew, 1978)

The propensity of a population to change will be the result of the interaction between the elements at the two levels, and so each may form a control dimension of the catastrophe model. The factors which operate at the level of society as a whole can be massed into a dimension which is termed the "ability" of the population per capita (A). This dimension includes more than just carrying capacity and economic variables such as welfare but also demographic factors, such as mortality rate, life expectancy and other factors like the educational level, attitudes to religion and concern over environmental issues.

The other control factor, labelled "willingness" to have children (and to allow immigration) (W) includes the worth of children in economic terms of the child's contribution versus cost in rearing, and in opportunity costs terms, where children are seen as a trade-off against increased leisure time, material consumption and parental activities such as travel and participation in paid employment. Also included in this dimension are the perceived need for children as risk insurance and as investments for the future, the effect of female emancipation, changing levels of divorce, adherence to fundamentalist ideas, innate desires to perpetuate the species.

The bifurcation comes between these factors when W is high. When this area is encountered with rising ability, the preventative check might dominate as children are traded off against parental consumption of time and manufactured goods and services. But as A continues to rise, anastrophe will result. If the area of bimodality is encountered as a result of decreasing ability, P will remain high until collapse, as a result of the positive check is inevitable.

This axial system is not independent but the control variables interact strongly with one another. The willingness to have children is dependent on the factors at the level of society such as welfare which, in turn, depend on the willingness to have children. This prevents conventional statistical modelling.

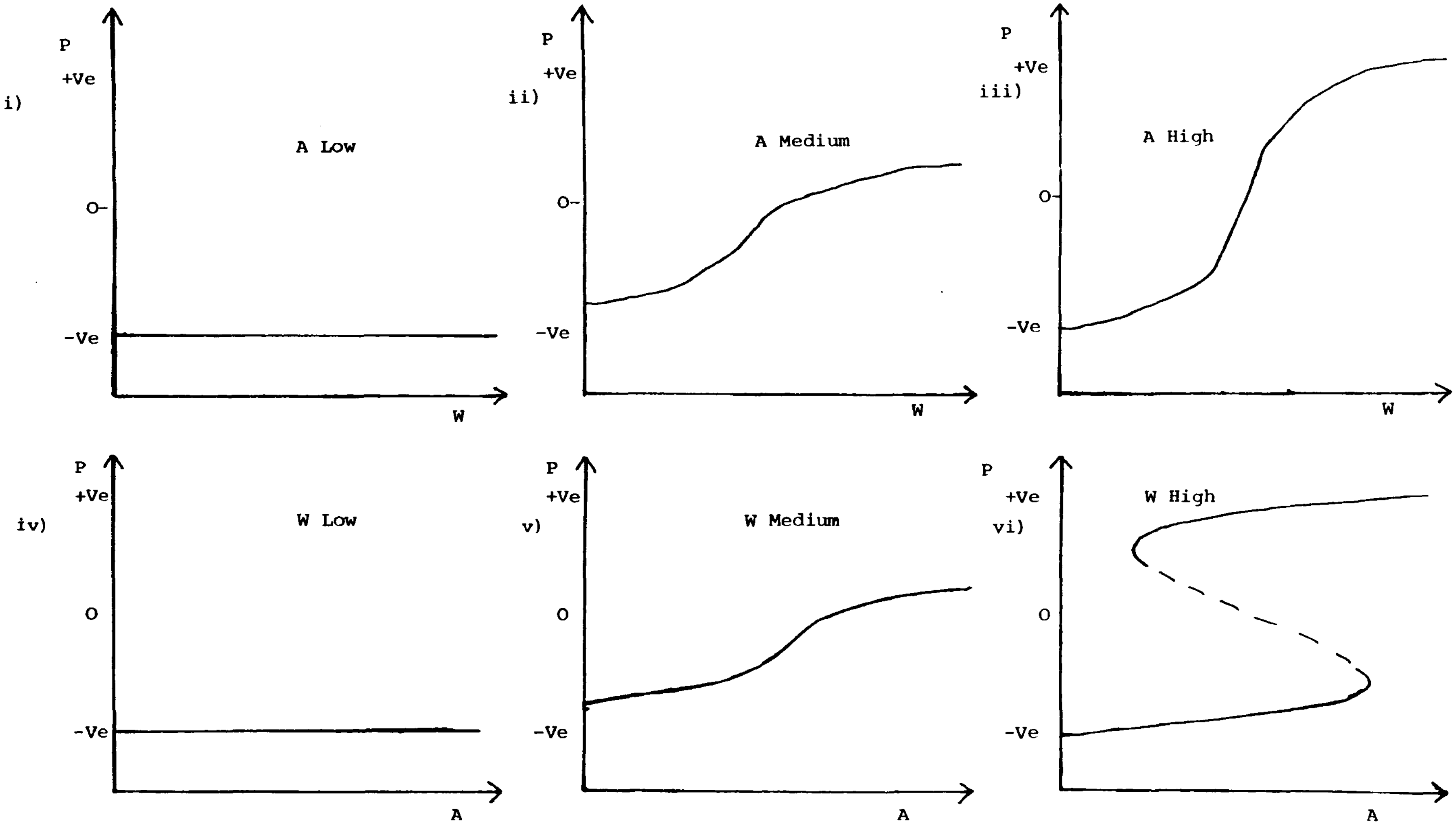
The argument is that the response surface is modelled as the function of the two controlling factors,

$$\text{i.e.} \quad P = f(A,W)$$

(P = propensity of the population to grow)

To visualise the surface, first consider the two-dimensional sections of the surface which are illustrated in Figure 6.4.3.

Figure 6.4.3 - The Extremes of the Population System Model



Note that in Figure 6.4.3(iii), when A is high, propensity to grow rises with W in a sigmoid fashion - this is to be expected as there is an extreme upper limit to the rate of growth, set by biological factors. In Figure 6.4.3(vi) the fold catastrophe is apparent - as soon as A becomes large enough and looks as though it has settled at the higher position, then there can be a jump to a higher level. When A falls, P remains high as W is high until transition must be made, and there is a sudden fall.

Combining the Figures in 6.4.3 into one surface gives Figure 6.4.4.

This surface models the propensity to grow for combinations of A and W and indicates the range of zero growth and that the cusp is only encountered from below when there is little or no growth and when W is high. From above the cusp is only encountered when A is falling and W is large. Some paths over the surface are shown in Figure 6.4.5.

Path 1 would be the path if the population was controlled by the Malthusian preventative check - it represents a negative feedback loop where, say, population is stationary, the economy improves (A increases) - this may also lead to an increase of W. Eventually the population uses up spare capacity and A will fall. In turn, W also falls and the population returns to a stationary state.

Path 2 starts as Path 1, but "moral restraint" emerges and children are traded off against desire to maintain status, and to participate more fully in purchasing manufactured output. Thus return to zero growth is made at an increased A by decreasing W.

Path 3 starts at the same position as Paths 1 and 2 but, as A falls, perhaps as a result of population pressure, W remains high. Thus the population encounters the pleat from above and collapses,

Figure 6.4.4 - The Catastrophe Model of the Population System

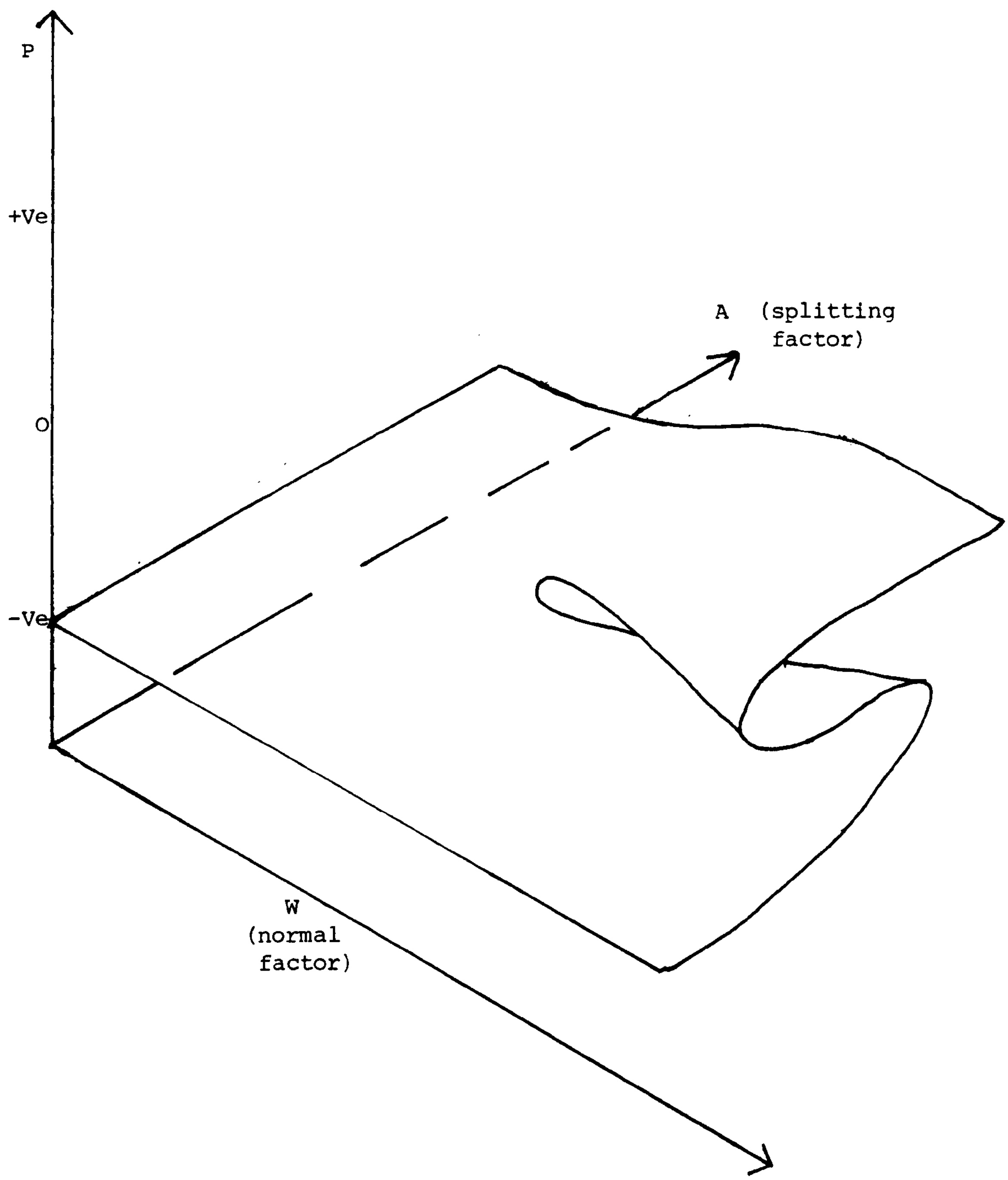
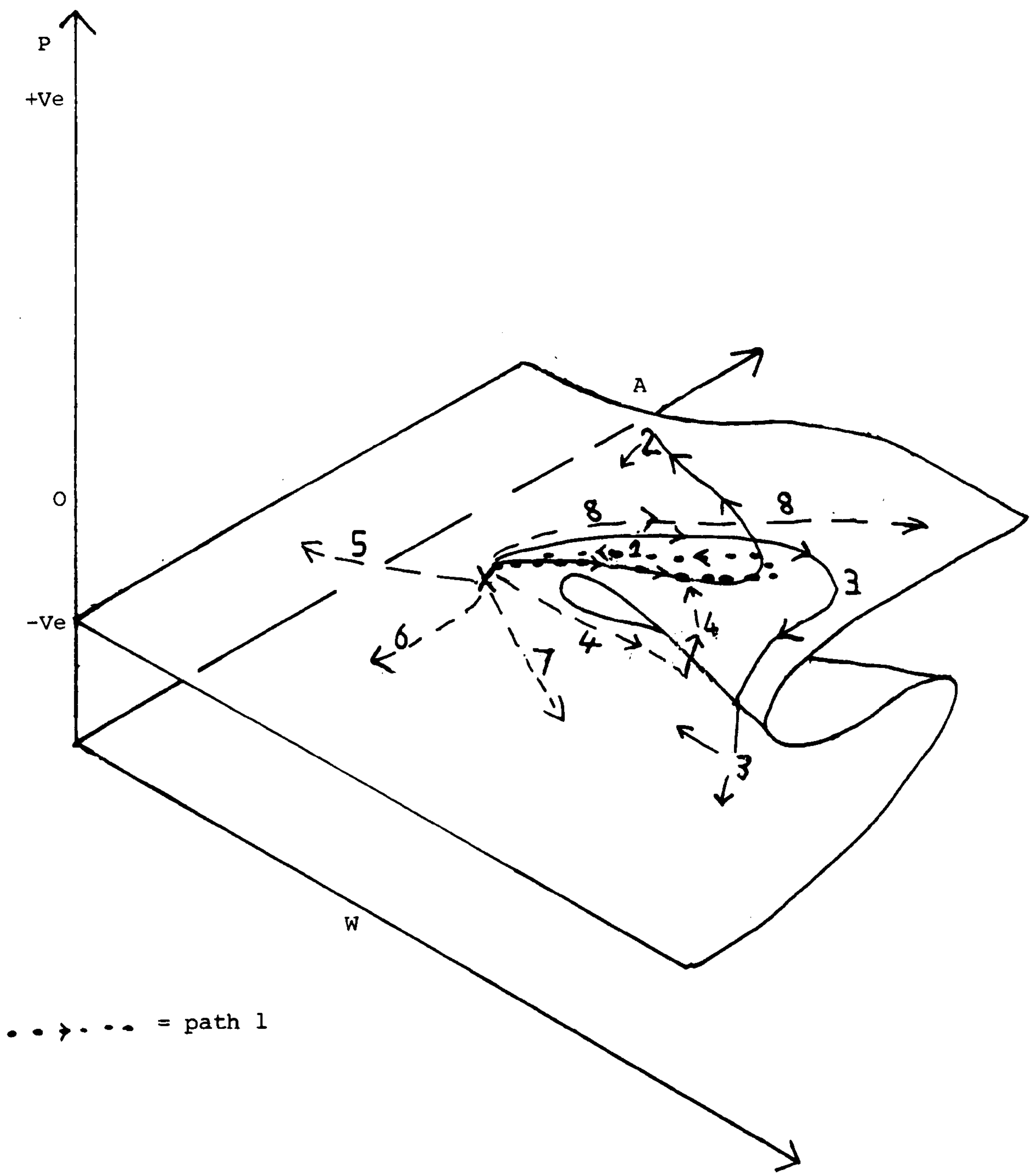


Figure 6.4.5 - Paths on the Cusp Catastrophe Surface



returning to zero growth or to prolonged decline if A keeps falling.

Path 4 starts as Path 1 but W rises faster than A, pushing the population under the pleat, and, if A continues to rise, anastrophe results. Return to zero growth can be made by any of Paths 1, 2 or 3.

Paths 5, 6 and 7 show continued population decline from a stationary position.

Path 8 shows continued population growth at a quickening pace - geometric growth.

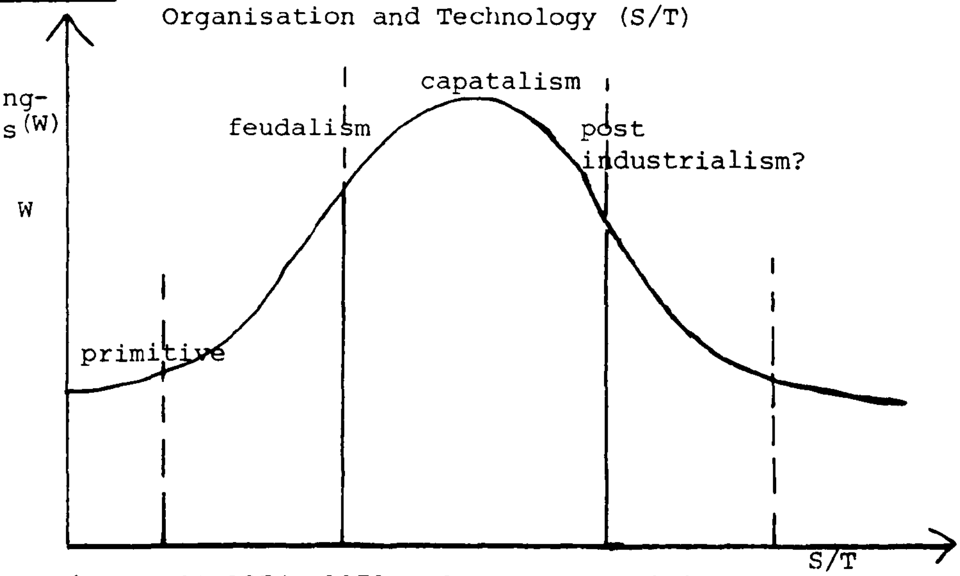
Paths 1 and 2 would manifest themselves as sigmoid growth curves. Path 3 would be a case of de-escalation, while Path 4 would give escalation. Paths 5, 6 and 7 could be viewed as negative exponential decay - if, however, these paths returned to zero growth, then the population could be modelled using an inverted sigmoid. Finally, Path 8 would appear as an exponential model.

This type of model forms the basis of a good description, in qualitative terms, of the population system, and gives a more comprehensible visualisation than block diagrams that are traditionally used by demographic system modellers. However, this model fails to account for major technological advances and changes in social organisation. How these might affect the A and W axes is illustrated in Figure 6.4.6. Figure 6.4.7 portrays this new model as a series of "snapshots" of the continually changing surfaces, as the society develops.

The effect of moving along this new social/technology dimension may be rather like an erosion process working on a landscape in that the high ground is eroded and deposited on the

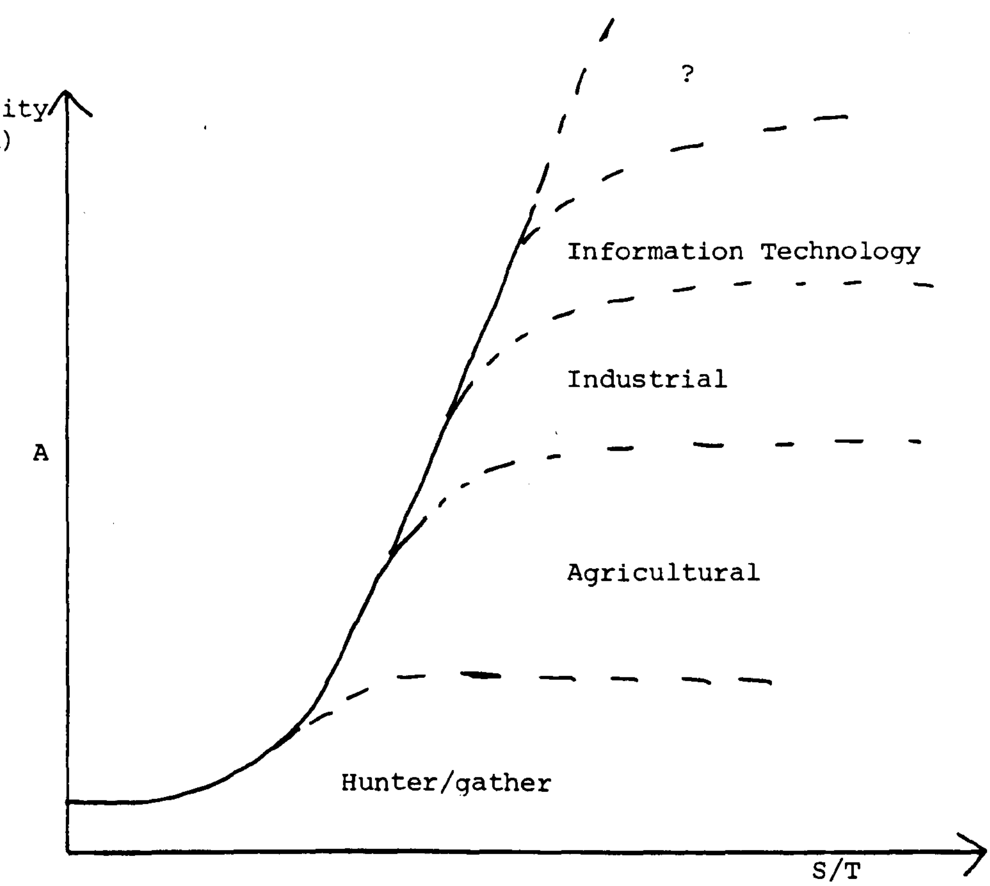
Figure 6.4.6 - Possible Effects of Changes in Social Organisation and Technology (S/T)

i) On Willingness Axis (W)



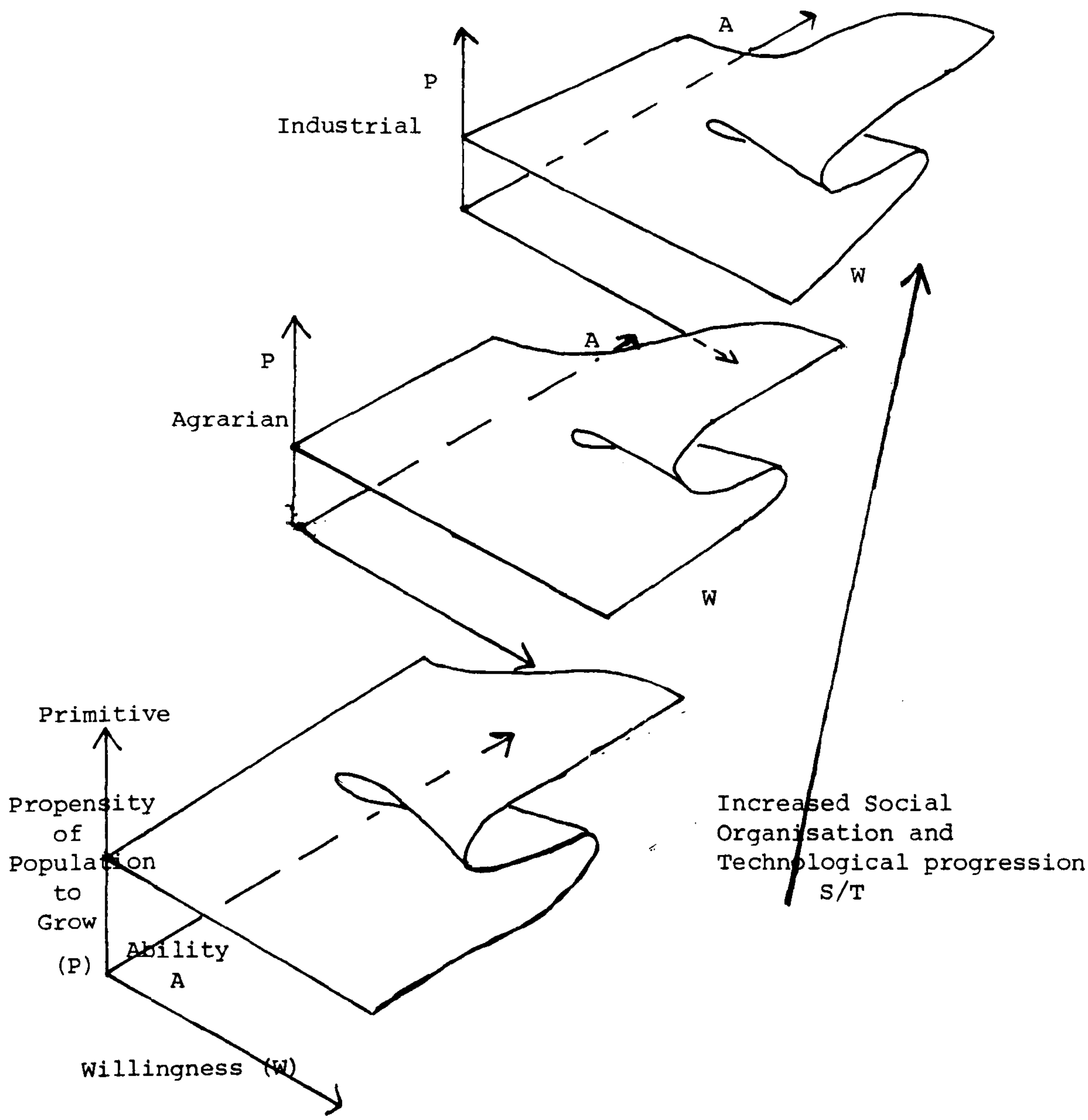
(See Bell 1974, 1979 and Gershuny, 1978)

ii) On Ability Axis (A)



In the type of process pictured above A increases as a succession of escaled Sigmoids (see Bell, 1974, Leach and Wagstaff 1986 and Marchetti, 1985)

Figure 6.4.7


P
Industrial
A
W
P
Agrarian
A
W
Primitive
Propensity
of
Population
to
Grow
(P)
Ability
A
Willingness (W)
Increased Social
Organisation and
Technological progression
S/T

low ground. Thus there is less likelihood of major depopulation or massive "explosions" of population growth rates, the more socially and technologically advanced the population becomes, as the possible anastrophes or catastrophes may become relatively less dramatic than they were in the past. This could be due to the accumulated ability and reserves of the population, both in capital and in knowledge terms, forming safeguards either to prevent or at least alleviate calamities, aid can be accessed from other regions, and the dangers of rapid population growth is recognised, and more sophisticated economic and social systems will possibly act as a braking mechanism.

Considering further the possibility for large jumps, they are perhaps diminished in likelihood with increased social organisation and increased technology, but they are still possible. For example, China's one child policy may well prove to result in a catastrophic jump in the propensity for growth. Other possibilities are that catastrophic falls might result from concern, which is highlighted by more pervasive media, over the AIDS virus, increased threats of nuclear war (or nuclear war itself, see Coale, 1985), economic problems resulting from energy shortages or environmental degradation, caused perhaps by the slow build-up of contaminates from increased use of pesticides, artificial fertilisers and pollutants. Catastrophic falls may also result indirectly for a developing country through the process of "Westernisation" (Caldwell, 1982) if there is an increase in social and technological organisation in the developed world which Caldwell argues will reduce fertility in the developing world. Social changes in the care for children, such as increased use of creches and child minders, as the "ability of the population" increases, may reduce the opportunity costs of children which might result in anastrophe in developed societies.

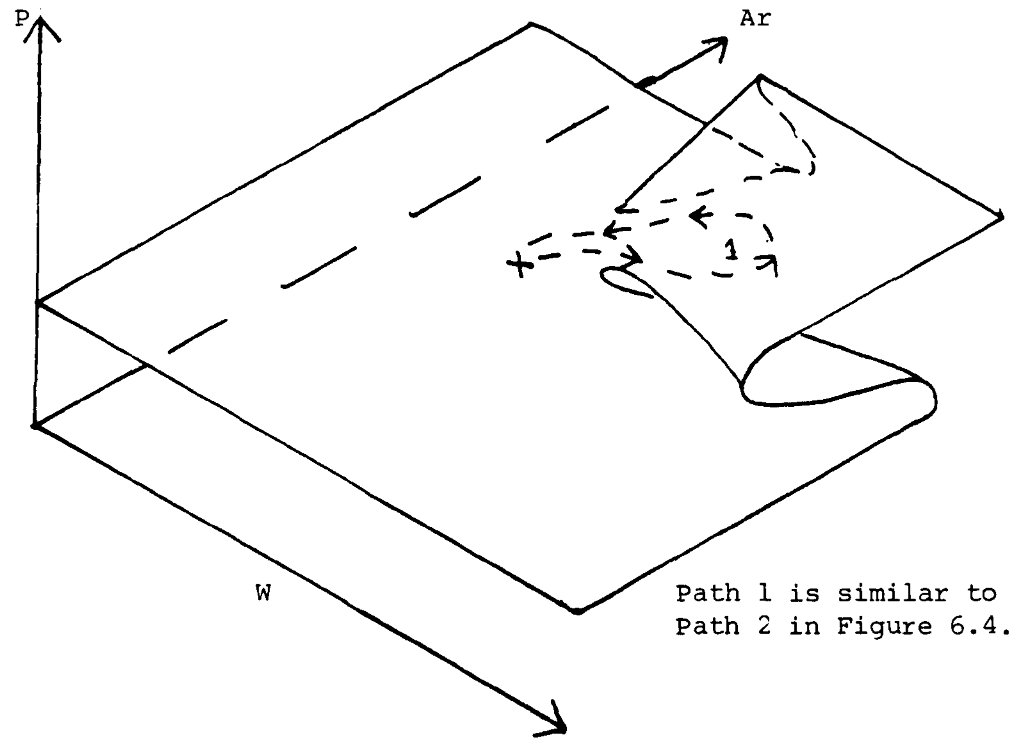
This then suggests that the further controlling factor of social and technological organisation needs to be incorporated into the model. From the discussion above, the effect of this is minimal until at fairly high levels of social and technological organisation when it would seem a pleat emerges perpendicular to the original cusp. This pleat is encountered with changes in willingness to have children.

To reflect the changes in the A dimension with increased social and technological organisation could be achieved if the A dimension is made the relative ability of the population per capita A_r. This scale would be relative to a past generation in perhaps some variant of the Easterlin (1968, 1978) approach to the study of the population system. Thus the ability dimension has become more fluid and can now incorporate changes in technology and in the manner by which society is organised. This new model is illustrated in Figure 6.4.8 for the case of high S/T (for low S/T the new model will be as in Figure 6.4.5). This model, which has three control dimensions, is a variation of Zeeman's (1976) Swallowtail catastrophe and is advanced as a general model to model the population system. However, since the pleat on the wellbeing dimension only becomes apparent with high S/T for most populations, in historical contexts the simpler cusp catastrophe can be used for most populations.

An important heuristic tool has now been developed; it is a model of population development that is of great generality. Positions on the catastrophe surface represent periods of when the population is stationary, growing or declining. It is argued in

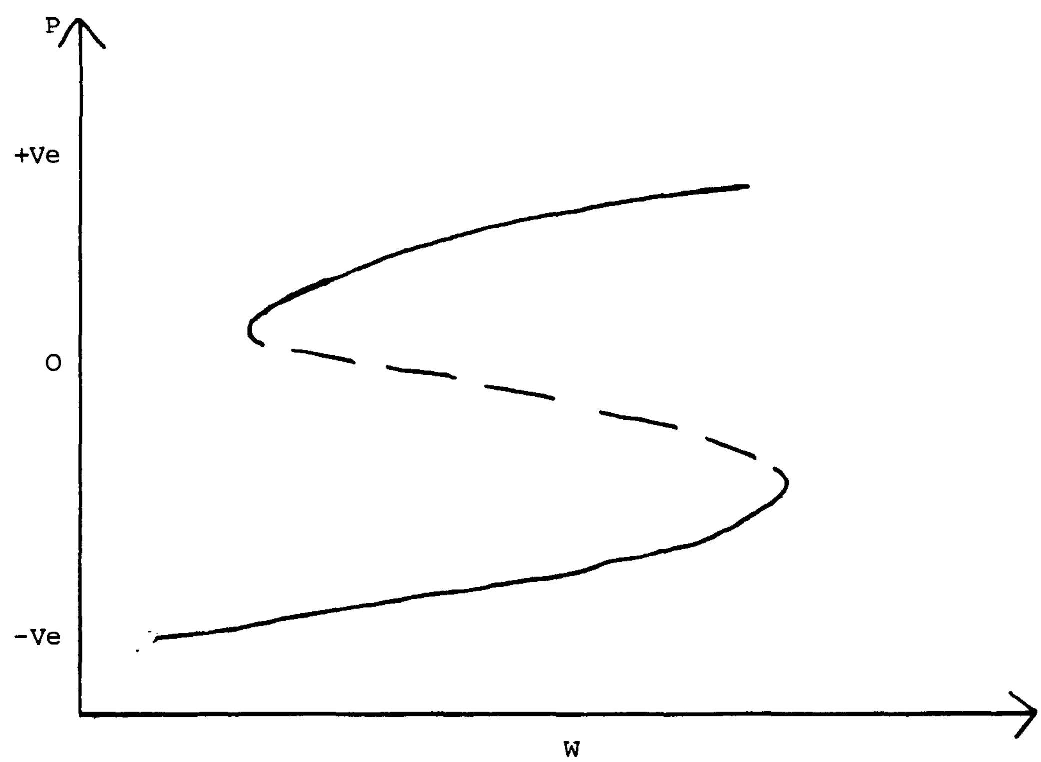
Figure 6.4.8 - Catastrophe Model of the Population System that Incorporates Changing Social Organisation and Technology

i) S/T High



Path 1 is similar to Path 2 in Figure 6.4.5

ii) S/T and Ar High



this thesis that, unless a region of bimodality is encountered, one should be confident in obtaining satisfactory forecasts from simple trend curves. When a region of bimodality is encountered, then, if movement persists in the same direction as phase change, a dramatic change in the pace of growth is to be expected. Hence one should greatly reduce forecast lead times.

Sketching the anticipated future movement on the surface explicitly states many of the assumptions, such as assumed economic growth, assumed changes in religious adherence, assumed attitudes to contraception and so forth, that are necessary to make forecasts. These assumptions are not stated verbally but are illustrated in an amalgam relative to the control axis. This means of summarising and displaying assumptions is not possible with conventional demographic models.

The use of the catastrophe model is demonstrated with reference to historical populations in the next section.

6.5 Application of the Catastrophe Model to describe Population Development

In this section, the model developed in the last section is used to describe some of the population changes which have been investigated in this thesis. The populations examined are those where one simple sigmoid was found to be unsuitable: Stone Age population, France, USA and Ireland; and Sweden and Great Britain where a simple sigmoid model was found to be satisfactory.

6.5.1 Pleistocene and Post-Pleistocene period and the Transition from Nomadic Hunter-Gathering to Sedentary Agrarian Lifestyles

As discussed in the previous Chapter, the change from hunter-gathering to pastoralism may well have led to a major increase in population size. This period may have been characterised by a change from a system in which there was zero or only slow growth, where a low density population was restrained by various constraints, both social and biological in nature (Cohen, 1980, Lee, 1980, Howell, 1986), into a system with a larger population living at higher densities (Hayden, 1981, Hassan, 1980). This change, Hayden (1981) suggests, may have followed a progression through the following states.

First, there was diversification of the "resource base" in that a wider range of animals and plants were seen as food and more r-species animals (May, 1981) were consumed. This could have resulted out of a growing scarcity of k-species animals which have long birth intervals, small litters, and long life spans, and developments in cooking and harvesting of small animals and plants. This increased co-operation between groups so that more settled social groups emerged in which children would be less burdensome. Thus the "ability of the population" increased. This led to restriction of the range of nomadic activity as there was less need to follow migratory herds of k-species animals such as mammoths. (This might also have been encouraged by climatic changes (Hassan, 1980)). From such a position the population had some of the checks on growth relaxed as carrying capacity rose. It is possible, for example, that the biological check of "fatness" (Lee, 1980, Howell, 1986) would have retreated and increased cereal consumption would have increased the potential for ovulation and, as there was less distance travelled, children, who were carried, would be less of an inconvenience. Children would also have higher economic value as they would be useful in catching and preparing r-species animals. Thus the "willingness to have children" would rise (the mechanisms being similar to those quoted by Caldwell (1982) in his wealth flow theories).

However, there are many explanations of population change at this time and it is a proposition of this thesis that (unlike Hayden, 1981) these changes were of insufficient magnitude to cause a major change in the population system. The changes were more likely a manifestation of population pressure in that they were a reaction to the diminution of the stock of k-species resources available as a result of over-hunting, and possibly climatic changes, lowering the carrying capacity. It would be unlikely that switching to r-species animals, when compared to the glamour and ease of hunting k-species, would appear to the population as an improvement in "ability". Also the switch to r-species would probably not give a more reliable resource base as these animals often follow "boom-burst" cycles as Maston (1981) points out. So the population may not have grown at all during this period.

A possibility is that the change in resource base was perhaps a preliminary stage which the population underwent and allowed attitudes to change to allow the settled cultivating of land and the domestication of animals. ('Settled' in this context does not mean permanent, and limited nomadism within an area would still be practised as, once the land used was exhausted, the population would move on to a new area in possibly a slash/burn agrarian system.) This would increase the "ability of the population" and would overcome the biological checks on ovulation and children would become economically worthwhile.

Gradually then the willingness scale would rise and, as more children would allow more land to be farmed, ability could also rise and a self-perpetuating cycle could emerge resulting in an anastrophic jump in the propensity to have children.

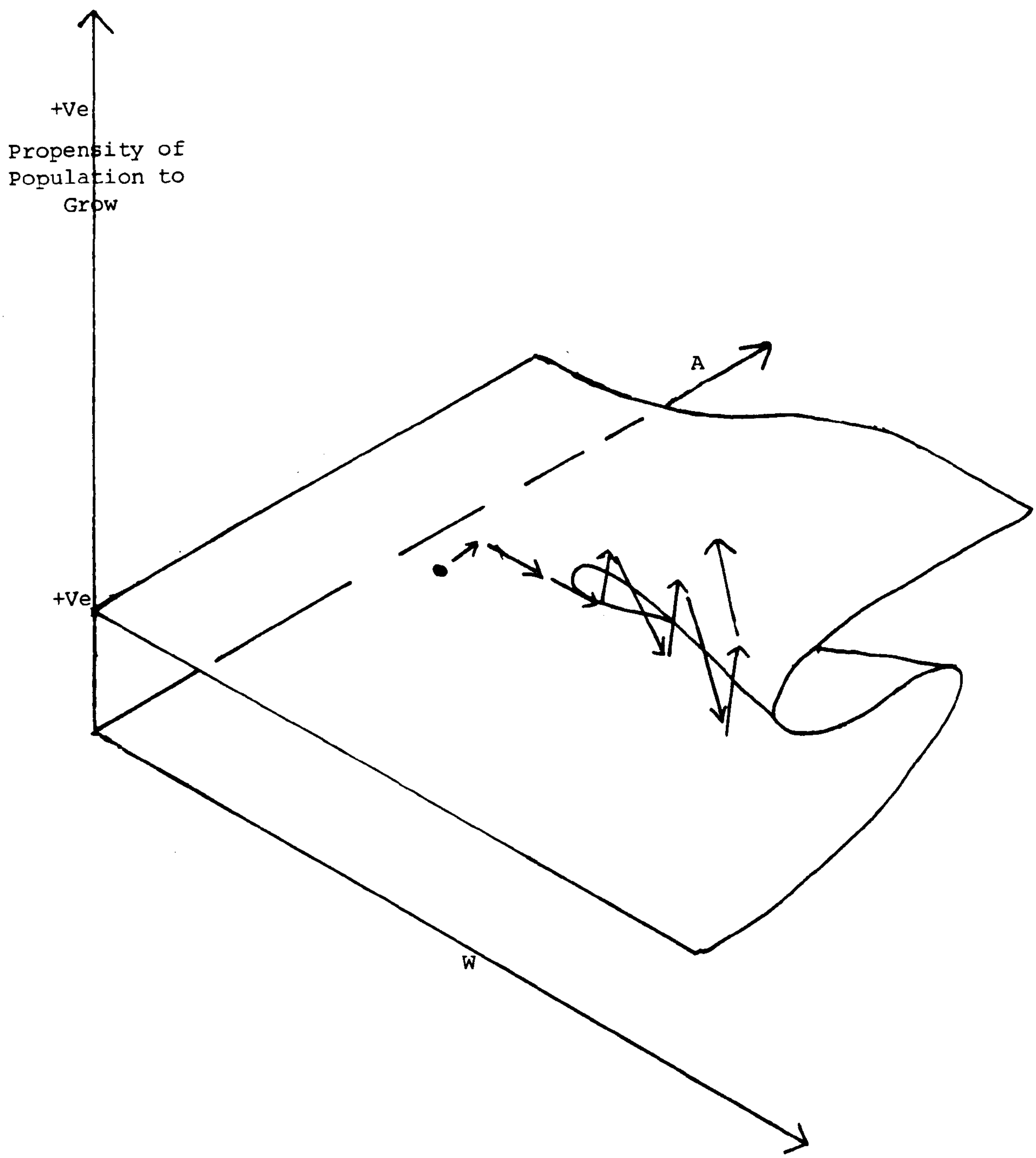
It is likely that this anastrophe, with the emergence of agriculture, led to a variant of a type of saw-tooth "boom-burst" behaviour in the system. This would develop because living at higher densities and in a more settled manner makes the population more vulnerable to increased mortality from infectious diseases and to malnutrition and starvation from crop failures. Thus "ability" would fall and, as the population were not yet very competent agriculturalists, there could be a drifting between the states of hunter-gathering and agriculture. However, although there would be an oscillation between increasing and decreasing "ability", it is conceivable that "willingness" to have children does not decrease but may well increase, pushing the population further along the pleat. Willingness rises as the benefits of having large families are apparent as a source of risk insurance, as a means of gaining and cultivating more land and in helping with arduous work. (Also Woods (1983) argues that fertility is directly proportional to mortality.) The population probably rose, not chiefly as a result of reduced mortality and increased fertility, but as a reduction of infanticide (Hassan, 1973).

During this phase the population gains more experience and competence in agriculture so that the frequency of catastrophic collapses becomes less and the population moves away from the cusp. (However, it is noted that this is dependent on the ecosystem and climate, changes in either could cause population collapse.) As a consequence of this improving ability, and increased skills in exploiting resources, larger social units may be formed and a state begins to develop in the manner suggested by Carneiro (1970).

The path on the catastrophe surface is illustrated in Figure 6.5.1.1.

The catastrophe model thus suggests that population growth came about relatively suddenly in these early times as a result of gradual improvements in ability and changing attitudes towards children. Therefore, it is in contrast to the "bottle" model which advances only one factor for the reason for growth, which is increased carrying capacity. Such an explanation would be unsatisfactory since there was little evidence of a major increase in carrying capacity, nor does such a model give any explanation as to why family size rose. The catastrophe model also suggests that there is more to the change in the system than the Boserupian explanation of population pressure fuelling innovation which, in

Figure 6.5.1.1 - Primative Catastrophe Map



turn, fuels more population growth.

6.5.2 Ireland 1750-1900

From 1780 to 1845 the population of Ireland grew from four million to just over eight million. But, by 1851, the population had collapsed to 6.55 million - some 800,000 to 1,000,000 died, mainly from infectious diseases (although 21,770 died from starvation) and more than a million emigrated.

Grigg (1980) saw this collapse as a consequence of increasing population pressure and the effectiveness of the Malthusian positive check. However, the Irish system requires more explanation than the population simply outgrowing its carrying capacity. For instance, there was little advance in the "ability" of the population, and this was partly due to the failure of manufacturing and industrialisation to evolve (perhaps being hindered by British restrictions on Irish industry and, even when these were lifted, Irish industry was in too much of a primitive state effectively to compete with British industry). The result was that the Irish population was kept mainly rural and dependent on agricultural systems. This was also encouraged by landowners, many of whom were British, who invested their money in British industry, although the Roman Catholic Church also benefited.

With the introduction of the potato as a response to high demand for animal fodder, and later for human consumption, a family could be supported on smaller plots of land. This encouraged subdivision of plots and earlier marriage. This was often promoted by the landowner in order to get more rent. Thus, by the late eighteenth century, Ireland had become one of the most densely populated countries in Europe, with 98 people per square kilometre.

Grigg (1980) points to signs of falling living standards, for example, there was a 14% increase in the landless population over the period 1831 to 1845, and average wages fell from £4 to £6 per annum in 1811 to £3.3 in 1821. Grigg also draws attention to the beginning of significant emigration from Ireland in this period. Despite this, the population continued growing and competition for land became severe, leading to a rise in rents and land prices (e.g. 1760 to 1815 - rent quadrupled). There were some demographic adjustments to these hard times and fewer were "prepared to marry on the tenuous security of a potato patch" (Grigg, 1980). But these adjustments were not enough, and there was little scope for rural urban migration as there was a shortage of urban centres.

Living standards received further blows from successions of poor harvests in 1765, 1770, 1795 and 1800 and harvest failures (1811, 1816, 1817, 1821, 1825, 1829, 1830, 1832-3, 1834, 1836, 1839, 1841, 1842 and 1844). In 1845 the potato blight struck and, in 1846, the crop was almost a total failure.

Thus the situation was one where the relative ability of the population was falling yet the population was still increasing. The preventative check failed to come into play effectively. This may have been because of the strong adherence to fundamental religious beliefs, and because the use of potatoes helped overcome household formation mechanisms and children had become a source of risk insurance (which was utilised).

As living standards continued to fall and population continued to increase, catastrophic collapse seems invevitable as the model given in Figure 6.5.2.1 clearly illustrates.

Thus this model has given a clear and concise description of how slowly changing variables of willingness and ability could result in the sudden collapse of the Irish population.

Grigg (1980) also pointed out that similar catastrophes occurred in Ireland in 1740-1741, Finland 1690s and Denmark 1650s. These might also be modelled by a map similar to the one in Figure 6.5.2.1.

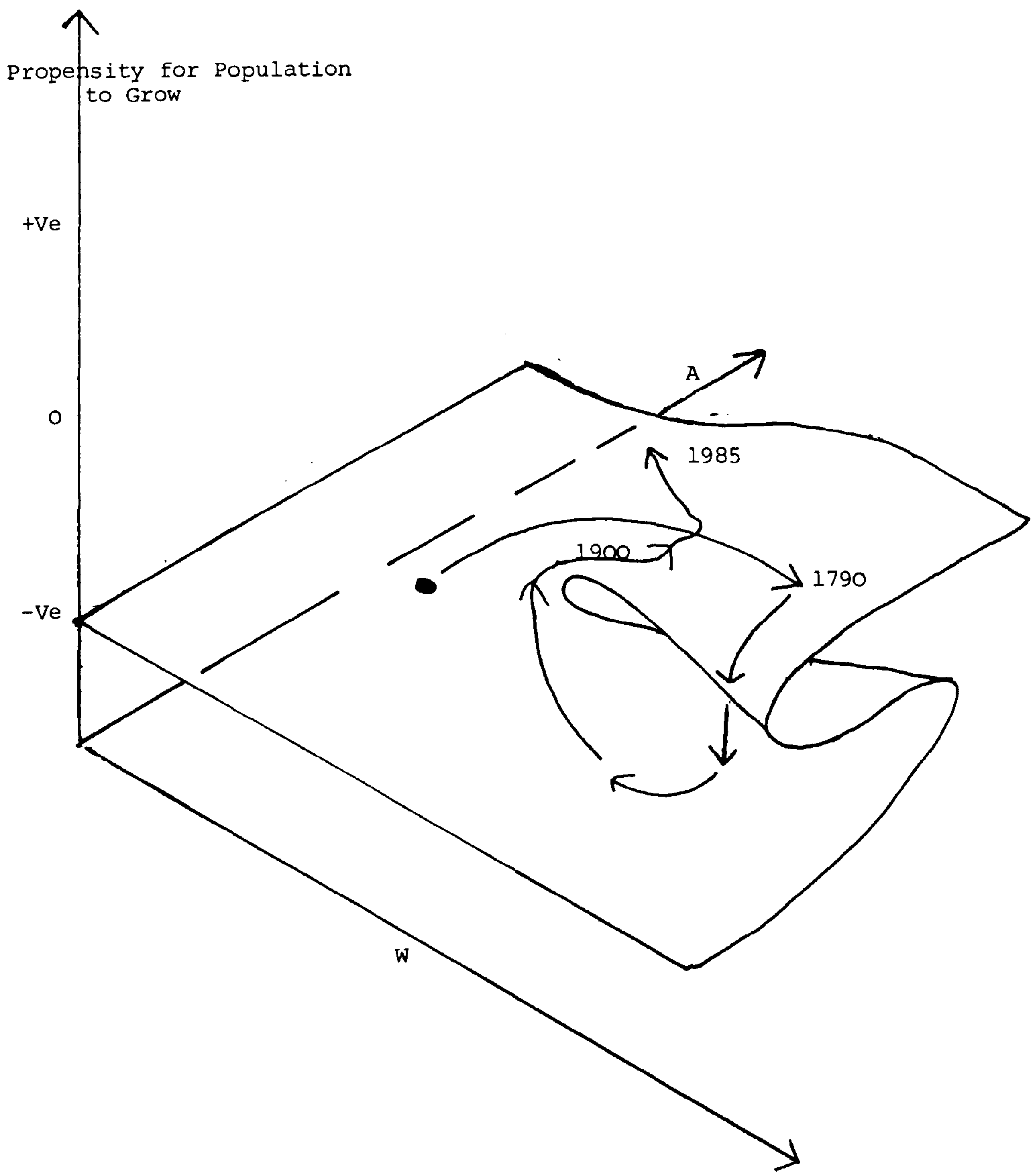
6.5.3 France 1800-1980

Over the last two centuries the development of the French population has been in contrast to its neighbouring populations. The French population grew at a far slower rate than other major European populations throughout the nineteenth century and became almost stationary by the early twentieth century. Ogden (1975) comments that by 1914 the French population was the slowest growing and had the oldest age structure in Europe.

This slow growth continued until the end of the Second World War - a three-parameter logistic model satisfactorily reflected this development. However, after the Second World War, there was a dramatic rise in fertility (the total fertility rate rose from 2.08 in the period 1935-1937 to 2.98 in the period 1946-1950). This rapidly increased the population size, and required escalated modelling. Post 1970, France is falling into line with other West European populations and fertility has fallen below replacement level (Huss, 1981). (The 1980 total fertility rate was 1.96.)

Wrigley (1985) discussed the dynamics of the first phase of growth. Mortality rates fell as in other European countries, but birth rates fell with a much shorter lag than neighbouring populations. This, Wrigley shows, was initially the result of delayed marriage and an increase in the proportion never marrying,

Figure 6.5.2.1 - The Catastrophe Surface for Ireland 1750-1980



but after 1820 there was earlier and more universal marriage and fertility control was practised within marriage. Wrigley (1985) writes of this behaviour that it was "a system of social regulation of overall fertility differing from the classic European pattern only in that it operated not just through the timing and incidence of marriage but also through the level of fertility within marriage".

It is considered that the practice of fertility control was a manifestation of population pressure (Wrigley (1985), Grigg (1980)). Grigg (1980) gives as evidence of population pressure the rise of rents, reduction of land in fallow, and the increase in number and fall of size of land holdings. Grigg states that, by 1881, for a large proportion of the population, the holdings were too small to provide minimum subsistence (the average size of holding having fallen by almost a hectare since 1851 to 3.5 ha). High rural population density was encouraged by the lack of urbanisation which was partly a consequence of the policy of subdividing land amongst heirs and the prevalence of small scale rural industry. This discouraged urban migration as occurred in the rest of North-West Europe (Birdsall, 1983). (Grigg (1980) illustrates the slow rate of industrialisation and urbanisation for, by the end of the eighteenth century, only around 20% lived in towns of more than 10,000 and, by 1851, this proportion had only risen to 25%.)

The different solution to population regulation might have resulted from an attitude change as a consequence of the French Revolution and a decline in respect for the clergy.

Hence, throughout the nineteenth and early twentieth centuries, ability and willingness were low.

By the 1930s, partly as a result of increasing ability, feelings of depopulation appeared within the population. This attitudinal change further developed in the Second World War and children became desirable and fertility rose. As a result, the population of France grew by 12 million over the period 1950 to 1980 (although immigration did make a contribution, especially from Algeria, see Calot (1981).)

These movements in fertility became issues of great public concern and became central in French politics. The population were very aware of demographic issues (see Huss, 1980).

Dryer (1978) suggests some reasons for this as:

(a) pro-natalist inventives and abortion and the propagation of contraceptive information and equipment was made illegal.

(b) Special wartime attraction of children, such as the procurement of additional rations and fathers of large families were less likely to be called upon or blamed for dangerous wartime activities.

(c) Dryer considered that many French people, after capitulation to German forces, considered pre-war France to have been underpopulated and attributed France's early lack of resistance to enemy forces to be due to a lack of young people.

(d) There was increased employment during the War (after capitulation) and immediately after the War in rebuilding France, so there was economic security.

(e) The cohort of the "boom" of babies, born after the First World War, were now reaching parentable age.

This suggests that there was a major change in attitudes relating to the willingness to have children during the Second World War.

When this was coupled with a rise in ability, when France was liberated, an anastrophic rise of the French population appears as inevitable. Thus, not surprisingly, the upper asymptote rose, according to the analysis in Chapter 2, from 42 million to around 58 million.

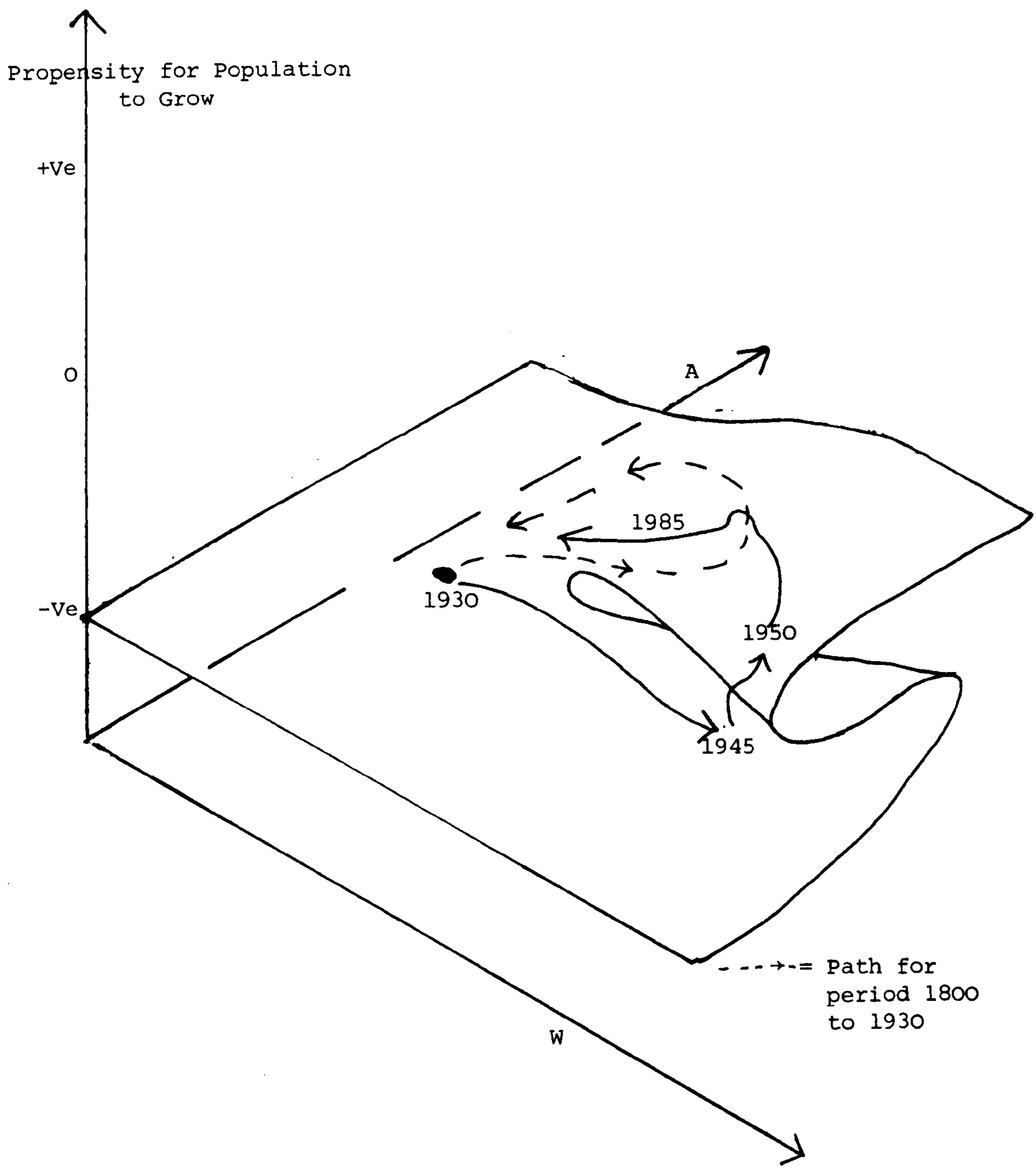
Recently though, French fertility is falling to similar levels as other European countries (Calot and Blayo, 1982), indicating that the escalated phase of growth is of short duration. This is a consequence of attitude change reducing the willingness to have children. Evidence of this attitude change was found by Girard and Roussel (1979) in that an increasing proportion of French women wanted a career job. In this situation they ideally wished to combine having children with a career. However, the general trend was to put careers first, even if this meant opting for less than their desired family size.

The catastrophe theory model of the French population is illustrated in Figure 6.5.3.1. By examination of the model displayed in Figure 6.5.3.1, one can conclude that escalation could have been predicted in the period 1930 to 1945 if opinion polls had been used to assess attitudes to fertility.

6.5.4 USA 1900-1980

Until the 1940s the population of the USA appeared to follow Pearl and Reed's (1920) logistic curve, approaching the upper asymptote of just under 200 million people. However, after the Second World War, growth rates increased as a result of increased fertility. (The total fertility rate increased by 46% from 2.523 in the period 1940 to 1944 to 3.69 in the period 1955 to 1959.) Throughout the period the "ability" of the USA slumped in the economic depression of the 1930s and increased as economic

Figure 6.5.3.1 - Catastrophe Surface for France 1800-1985



welfare, as measured by GNP, increased at a faster rate after the Second World War.

As ability rose, the trend to zero population growth continued as there appeared to have been a change in the willingness to have children. This may be indicated by the rise of female labour participation, increasing opportunity cost of children in that more consumer items were available, child education costs rose and employment opportunities for children diminished.

During the Second World War American industry was stimulated by high European demands as well as internal demand to equip the war effort, and the ability of the population started to rise.

This rise in ability may have set in motion a growth of suburban house building, which increased by over threefold, after the Second World War. This prompted a drift outwards from city centres. This improved housing and more affluent times led to a rise on the "willingness to have children" scale. "Willingness" also rose because of a partial ending of the old marriage pattern with more people marrying and marrying earlier (see Glass, 1967). The perception of increased instability of marriage with rising divorce rates might also have acted to reduce fertility. (Although the effect of divorce was small after the Second World War, the divorce rate was almost double the rate before the War.) Another factor affecting "willingness" may have been the fall in maternal mortality rates from more than 60 per 10,000 prior to 1935 to less than 10 per 10,000 after 1950).

Some reference must also be made to influences identified by Easterlin (1968, 1980). This relates to the catastrophe theory methodology in that both ability and willingness to have children rise with a small cohort of parentable age. This was the case before the Second World War, partly as a result of their parents being of a large cohort with low ability and willingness. The effect of this was manifested by earlier marriage and higher fertility when this cohort reached parentable age in the late 1950s and early 1960s.

The small cohort, coupled with improving economic conditions, meant that the USA population was "pushed" under the cusp. Further rises in ability (possibly coinciding with a peak in economic long waves (Rostow, 1960)), manifested by improved housing and economic prosperity, is sufficient to cause the population to undergo an anastrophic jump to the upperside of the cusp in the mid 1950s.

However, since then, with the growth of "consumerism" (Bell, 1979), increased emancipation of women, increased travel and involvement in active leisure pursuits created more "trade-offs" for children and willingness to have children has fallen. (The 1979 total fertility rate is down to 1.81.)

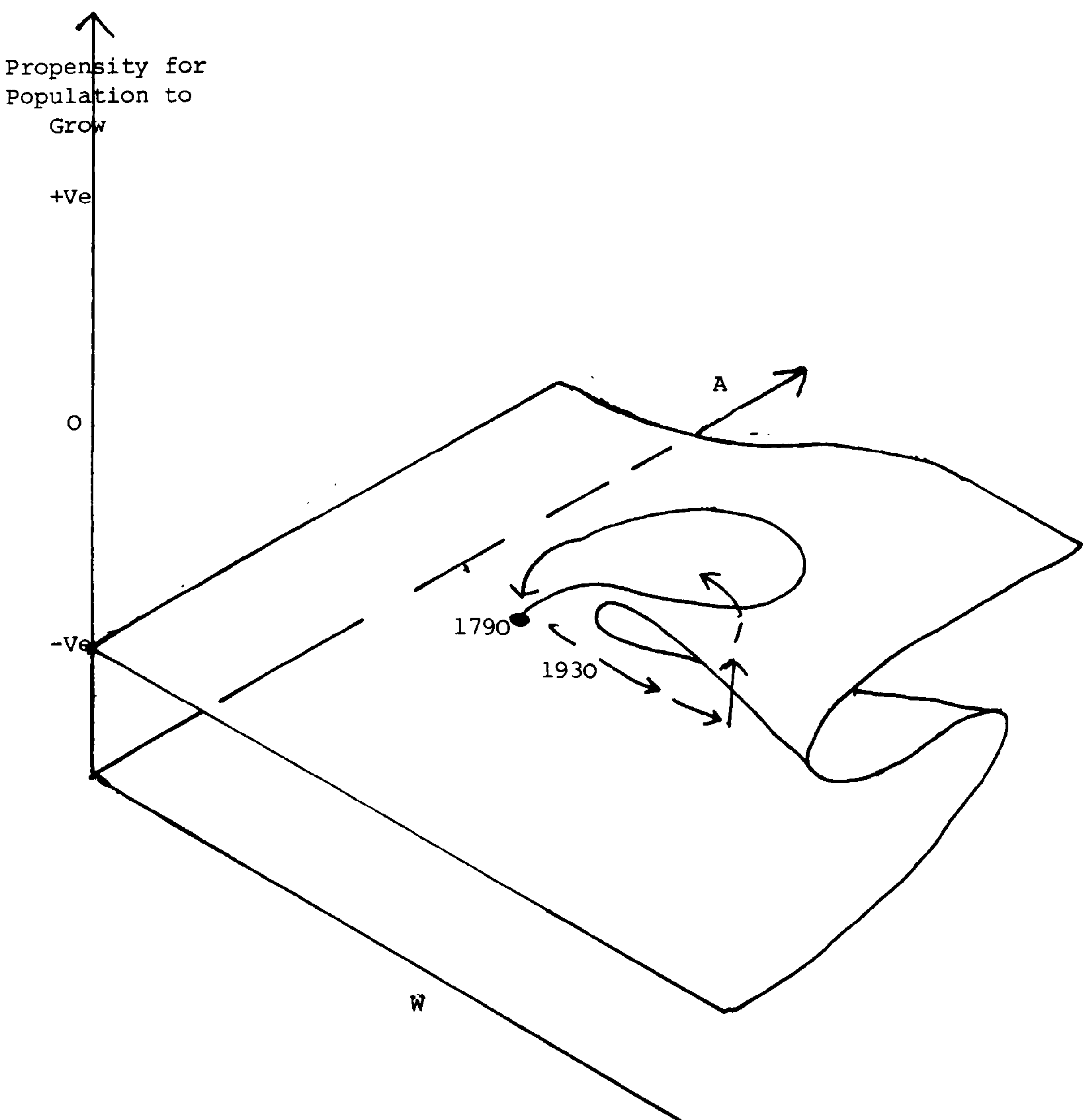
Since the cusp is encountered and passed at fairly low levels, as the population had not moved far under the cusp before improved ability led to escalation, the renewed growth phase, like France, may not last particularly long. From the results given in Chapter 2, the upper asymptote for this new growth phase is likely to be under 265 million. (This will probably be overshot due to ageing of the population.)

The path of the USA population is given by the model in Figure 6.5.4.1.

6.5.5 <u>England, Great Britain and Sweden 1800-1980</u>

These populations are examples where one sigmoid satisfactorily modelled the population as has been demonstrated

Figure 6.5.4.1 - Catastrophe Surface of Growth of U.S.A. Population 1790 to 1985

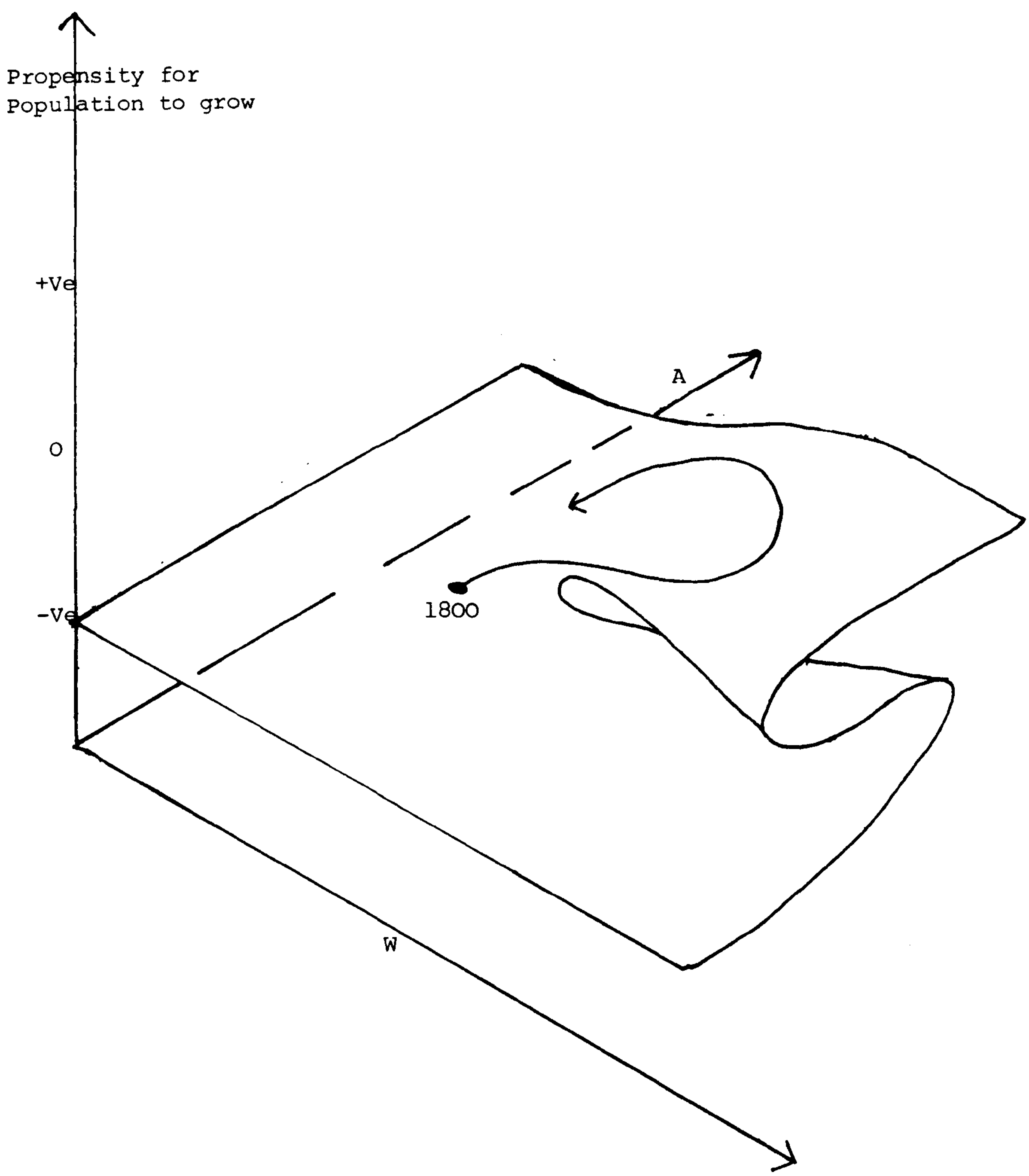


Path is the path described by the three parameter logistic model for the period 1790 to 1930

Path - → - is the 1930 to current movements of population growth in Chapter 2.

In these nations ability rose throughout the nineteenth and early twentieth centuries as a result of agricultural improvements, industrialisation, improved transport and trade, community medicine and sanitation schemes. In the nineteenth century the willingness to have children also increased, as there were little opportunity costs for children, but as urbanisation developed and parents became interested in purchasing manufactured goods and participating in leisure activities, willingness to have children has fallen. Also costs of children rise as it becomes difficult for them to participate in work that becomes increasingly non-familial, their education costs rise and children compete with leisure activities and consumer goods for their parents' income and time. Hence the preventative check came into operation. Woods (1983) clearly illustrates that, as standards of living rise above a certain level, fertility falls and falls to a socially defined minimum level. Also, women are more economically independent, there are heightened fears of divorce (Greene and Questor, 1982), and religious doctrines cease to be adhered to fundamentally (Simons, 1986). Thus the willingness to have children falls and there is a slowing of population growth. Figure 6.55.1 illustrates the model for these populations where no major disruptions to sigmoidal growth have occurred.

Figure 6.5.5.1 - Populations where a Single Sigmoid Model is Satisfactory



6.6 The Catastrophe Models reconsidered

In section 6.4 a generalised catastrophe theory model of population was developed. This is the Swallowtail catastrophe. However, in the previous section, several different population histories were successfully and succinctly described, using the simpler cusp model.

Thus, in the medium term, the simpler model should be used. It is only in the long term, when major social and technological change may be encountered, that the complex swallowtail model will be useful.

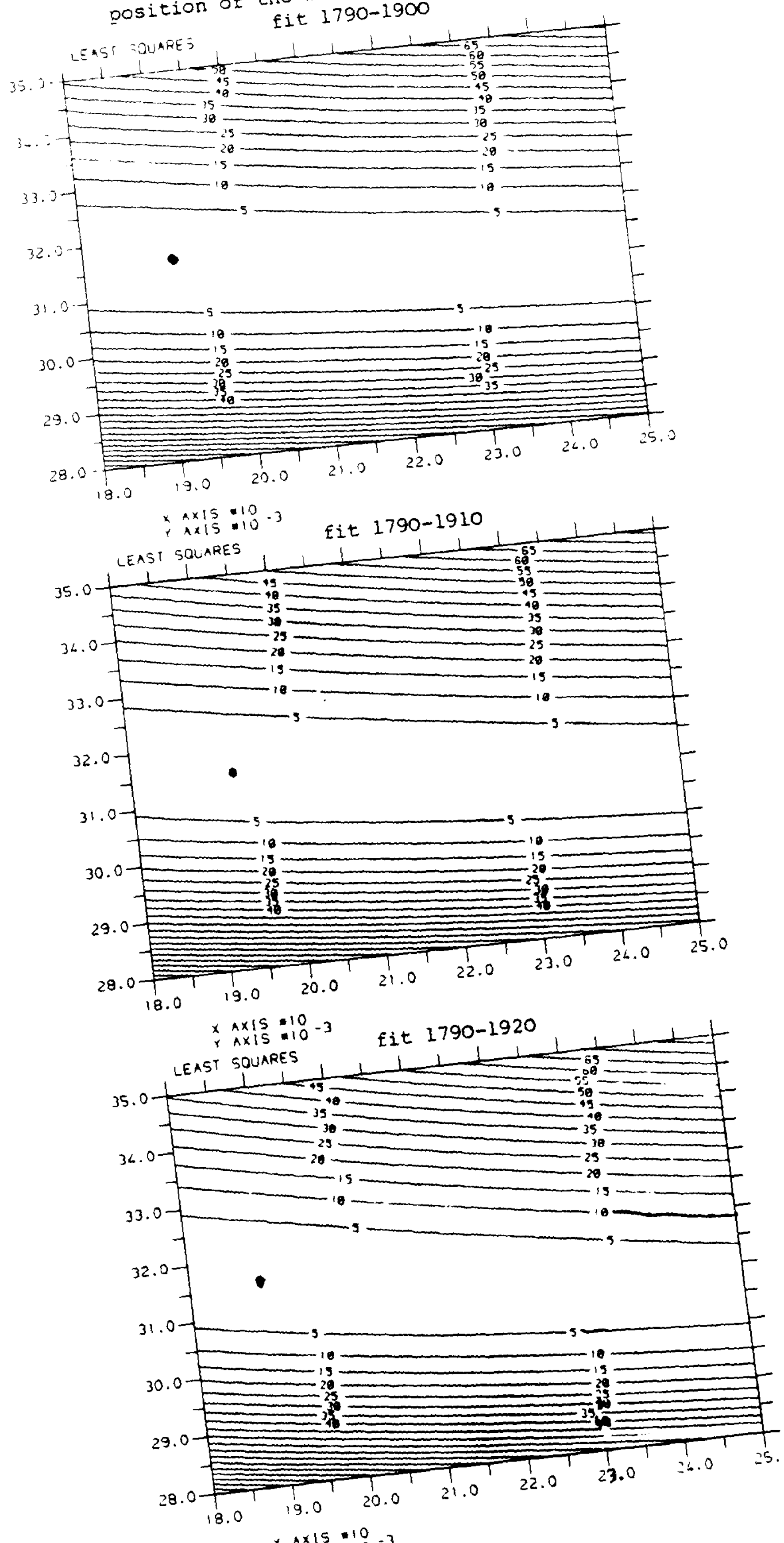
The model which gives a simple and yet a holistic description of population growth within the complex population system aids the study of population development and proves particularly beneficial for giving a framework in which assumptions can be clearly and parsimoniously stated. Such qualitative modelling is of use in its own right in aiding the formulation of traditional demographic models and would allow one to assess, as has been described in this Chapter, the stability and suitability of quantitative models.

By postulating a population on a catastrophe theory surface, one can also predict the likely future states of population growth by considering the future possibilities for the control factors assuming, of course, that the surface developed is suitable for future states of the system. This has been demonstrated in the previous section to be particularly useful in detecting the onset of escalation or de-escalation.

Empirically to validate these catastrophe models by quantitative means is not possible and inappropriate for reasons similar to those given in section 6.2 to dismiss the use of

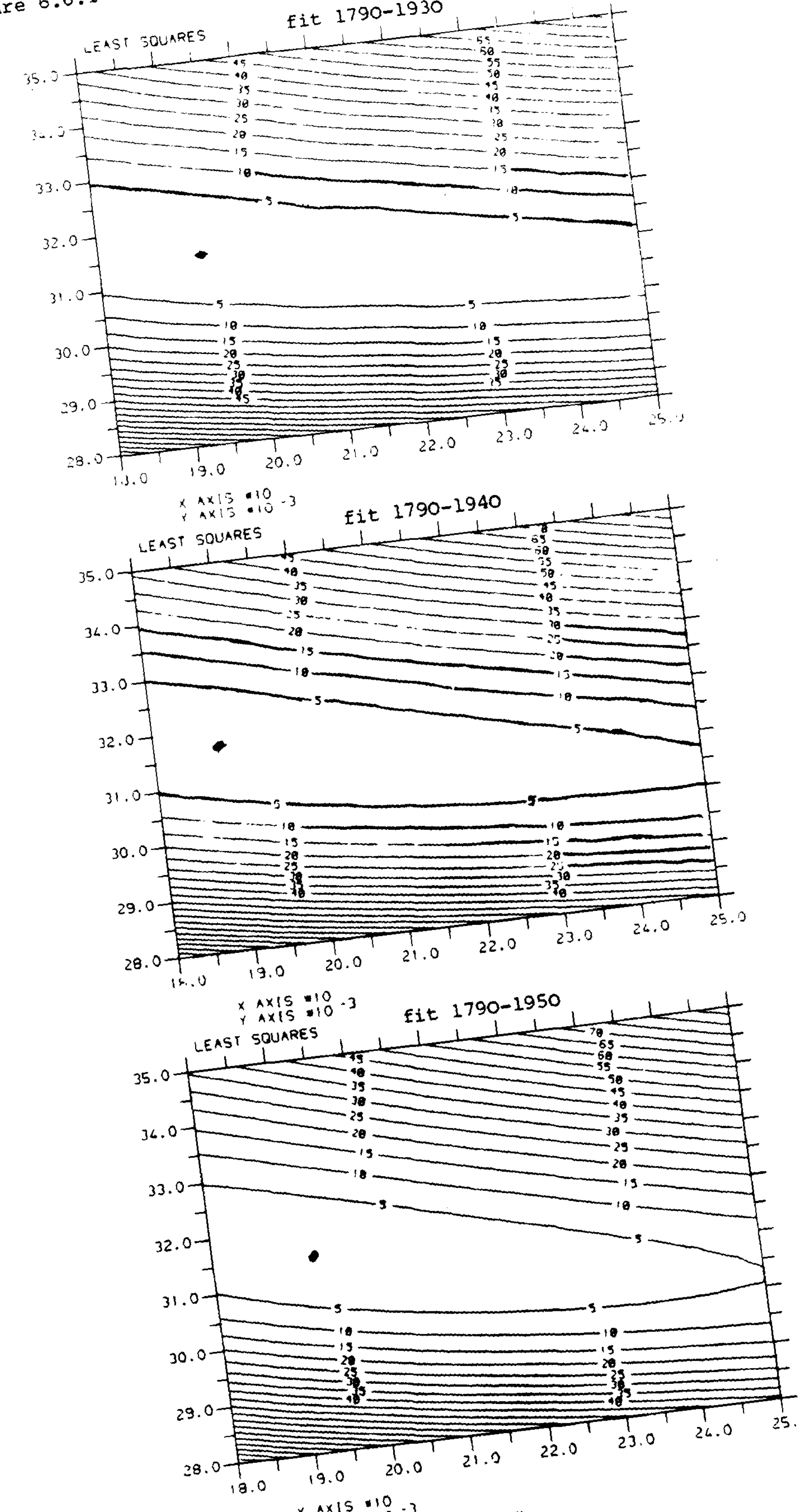
quantitative modelling. However, when contour surfaces of successive least squares fit, where compared, they do seem to confirm the presence of bimodality, sudden change and inaccessibility. This is illustrated in Figure 6.6.1 for the population of the U.S.A. which shows the position of the minimum on successive contour surfaces resulting from least squares fit (the N_0 parameter is held constant). The vertical scale of the contour surfaces is shown as a percentage of the sum of squares value at the minimum point. In the next section, the use of catastrophe theory models
in predicting future states of current populations will be investigated, and use will be made of catastrophe theory methodology to assess the appropriateness of quantitative forecasts made from sigmoid curves.

Figure 6.6.1 Contour surfaces from least squares fits to the USA population indicating a sudden change in the position of the minimum



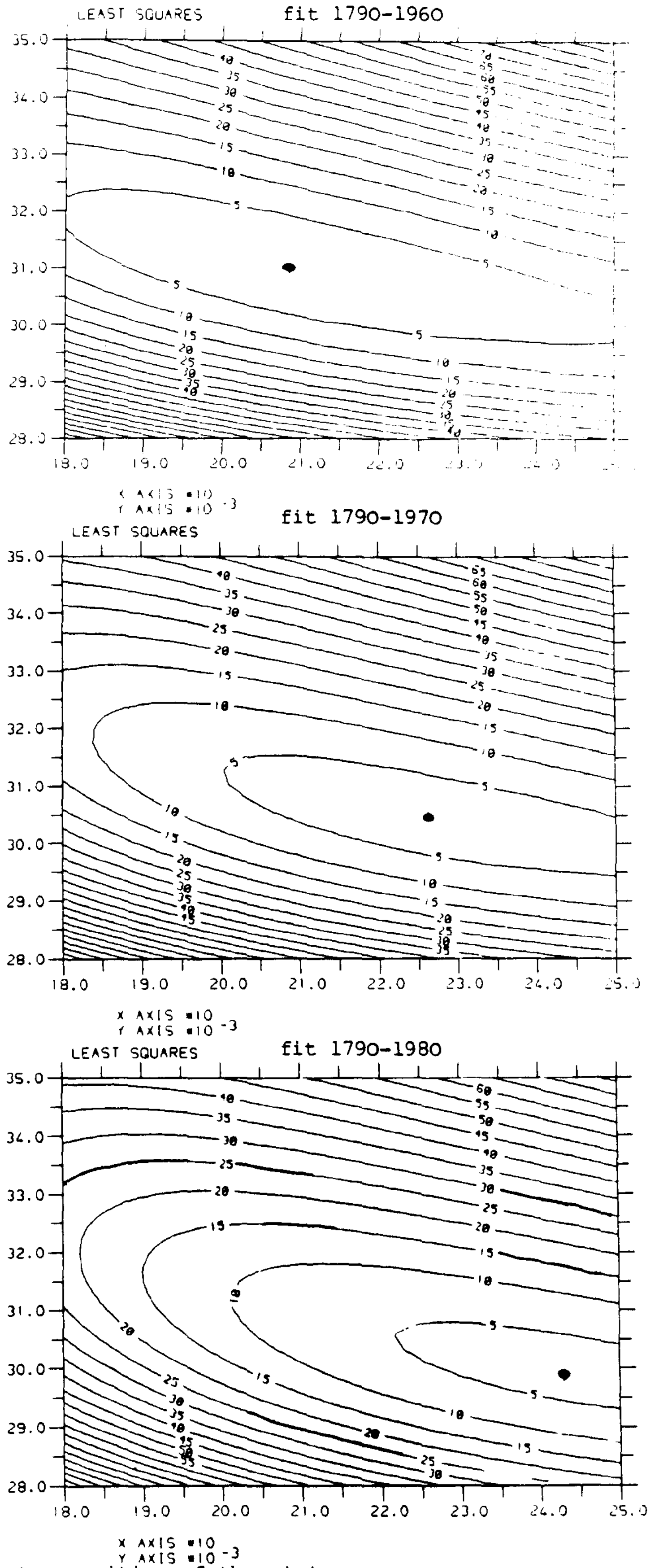
• = position of the minimum

Figure 6.6.1 - continued



● = position of the minimum

Figure 6.6.1 continued



• = position of the minimum

6.7 Future States of the System

By considering in qualitative terms the likely futures of the control factors, the probable paths of changes of population growth can be followed on the catastrophe map. Doing this will aid a forecaster to determine safe lead times for extrapolation from the current model of population size and to speculate qualitatively the degree of confidence in predictions.

In this section the future of the currently advanced technological/industrial nations will be considered first, followed by an investigation of the developing countries. The section is completed by an account of a catastrophe theory model to assess the future population trends of the less developed nations.

6.7.1. Advanced Nations

The models fitted to advanced nations in Chapter 2 and the forecasts made in Chapter 3 (and other forecasters such as Keyfitz, 1979, and Westoff, 1983) all indicate the population growth rates will tend towards zero. When the future of population growth in the advanced nations is considered by using the catastrophe theory methodology developed in this Chapter, one arrives at similar conclusions as it would seem that the ability dimension will remain high and the willingness to have children will remain low.

By extrapolating current trends no significant increase in willingness is to be anticipated as the following are likely to increase:

(i) economic and cultural independence of marriage;

(ii) less need for children for support in old age and as risk insurance;

(iii) little opportunity for children to contribute economically to the family;

(iv) later marriage and increased insecurity of marriage (see Murphy, 1985);

(v) decline of fundamental religious values (Simons, 1986);

(vi) higher opportunity costs of children as participation in leisure activities and the abundance of consumer alternatives increase.

This suggests that there will be a continuation of the trend to low family sizes.

Regarding the ability dimension, provided that there are no major changes in the economic status quo between nations, ability seems unlikely to decrease. Individual measures of ability appear to be rising, such as house ownership, amount of leisure time, and improved health, nutrition and life expectancy. An exception to increased ability might be in social-psychological wellbeing as society becomes more complex, but there is little evidence of this (Cohen, 1980).

It should be noted though, that improvements in ability have not been uniform across the developed nations, for instance, ability at the individual level in the U.S.S.R. could be falling as is highlighted in falls in life expectancy over recent years.

Some commentators such as Gershunny (1978), Freeman (1982) and Jenkins and Sherman (1979) predict gloomy futures for ability in the developed nations as they lose competitive advantages over emerging nations and refer to new technologies causing the "Collapse of Work" due to technological unemployment.

Bravermann (1974) refers to increasing alienation arising as a consequence of the application of the workplace of larger amounts of capital equipment.

Such a depression in the ability dimension might result in massive economic and social disruption. According to writers such as Bell (1974, 1979), this may result in capitalism failing to be an appropriate form of social organisation. Bell (1974) suggests that there will be a change from a society organised around "economising mode" of production and a sophisticated, responsible post-industrial society governed by technological elites. Others such as Schumarker (1973) advocate a return to a smaller, more primitive self-sufficient society.

Even with these radical changes that might have an effect on ability, it is unlikely that in a period of up to fifty years effects will be manifested in population growth, since willingness may not change and the time scale of changes is outwith the concern of this thesis.

Also, other writers such as Armstrong (1985) criticise and find no evidence of the notion of technological unemployment. It is considered that demand saturation is unlikely when there is such a large potential market if the less developed countries could be transformed to potential customers. Problems which emerge are merely those of "adjustment" which can be overcome.

There are other possibilities of future collapse of population which are present, such as a direct result of nuclear war or indirectly as a consequence of a nuclear winter. Coale (1985) gives some possible scenarios. Another possible cause of population collapse is of climatic change, possibly as a result of pollution (e.g. Broeker et al., 1985, Austin and Brewer, 1971).

However, these are not necessarily the result of variables identified in the system and their consequences are too unpredictable to be represented on the catastrophe surface. Thus they are noted as possiblities and are not pursued further.

6.7.2 Emerging Nations

Sigmoid models were found in Chapter 2 successfully to describe past growth of emerging nations represented by nations in South America, Mexico and India. Prediction using these models shows that one should anticipate falling population growth rates among such nations.

By using the framework of catastrophe theory one can argue that "willingness" to have children will fall and the "ability" of the population will rise (or at least maintain relative position).

The factors that influence willingness are likely to fall as there is:

(i) increasing development of industrial bases in such nations and declining familial based production;

(ii) increasing education;

(iii) establishment of rudimentary community and institutional care;

(iv) increasing urbanisation;

(v) several emerging nations are pursuing vigorous anti-natalist policies, such as the Chinese Peoples Republic and India (Kirk (1984) discusses the possible effectiveness of these policies).

However, the decline in "willingness" is unlikely to be uniform across these "nations". In Muslim countries strong adherence to fundamental religion, with subordination of women's role, is likely to continue and to delay significant falls in fertility. Also in Brazil, Daly (1985) points out that the process of fertility transition is hampered by controlling elites in Brazilian society who promote high fertility amongst the poor social groups due to economic advantages they accrue (such as a large pool of cheap labour).

The ability dimension seems to be rising as is indicated by increasing life expectancy and economic growth encouraged by technology transfers from the developed nations.

However, negative features on the ability dimension are also apparent, notably in Latin America where rapid urbanisation has raised considerable problems of social and psychological adjustment, and there are severe environmental threats emerging as a result of high levels of pollution and deforestation (particularly in Brazil).

Leontief (1983) warns that the future growth of ability for emerging nations is by no means inevitable. With an innovation of increasing automated production, the competitive advantage over developed nations of emerging nations of cheap labour will be lost. Leontief cites an example of a Japanese spinning mill (low grade textiles would have seemed to have been an ideal candidate for transfer to less developed countries) where, at most, only ten people per shift are required to produce sizeable output (30,000 ring spindles). However, it is likely that "following" nations will always have competitive advantages as set-up and development costs are less assuming a certain base infrastructure is obtained. Increased automation also overcomes a problem for emerging industrial populations in that there is less demand for a skilled well-educated workforce. So, on the whole, national economic ability of emerging nations is anticipated to rise. Hopefully this can be spread throughout the population by the provision of public works, such as road, school and hospital construction.

Overall though, for the next half century there seems little reason to anticipate that the populations of the emerging nations will encounter a cusp, and so one should have confidence in forecasts made by sigmoid models.

There are possible exceptions to the above, for example:

(a) Nations such as Mexico with high economic debt, international financial crisis, could herald severe restriction of lending to such countries resulting in economic collapse and possibly a catastrophic fall in population.

(b) The case of the Chinese Peoples' Republic (C.P.R.) is of interest. Currently the leadership of the C.P.R. are imposing strict anti-natalist policies (for details, see Chen, 1982), in the hope of curbing population growth. This could, when viewed in the context of the swallowtail catastrophe, by reducing the willingness to have children, cause the C.P.R.'s population to collapse by "falling" over the cusp at right angles to the willingness dimension. Since the level of social and technological development is unlikely to have reached a sufficient degree to allow full development of this cusp, the fall in population will be slight. From this position, alternative scenarios are possible, for example:

(i) if ability rises, there could be a backlash against anti-natalist policies perhaps seeing them as unnecessary. This could lead to an anastrophic jump to a much higher growth rate, as the cusp would be more developed, or

(ii) if ability falls (possible as a reaction to the anti-natalist policies) and a backlash against the anti-natalist policy emerges, increasing willingness, and then ability started to rise, anastrophe could be encountered.

These possiblities for the C.P.R., illustrated in Figure 6.7.2.1, show that, by restricting fertility, with a view to rising ability may lead to a far higher population in the long run.

6.7.3 The Underdeveloped Nations

The poorest nations of the world, such as most of the Central African countries, are the nations that are exhibiting the fastest growth rates (the high "fertility belt" (Lesthaeghe (1980)) and the population of the developing world seems to be on a path that would result in a tripling, in size, in the next century (Gwatkin and Brandel (1982)).

The relative ability dimension has an enormous scope for improvement but it is unlikely that economic welfare will improve by much over the next fifty years. The reasons for the poor prospects for development include:

(i) poor agricultural land;

(ii) lack of demand for industrial products;

(iii) lack of political and economic experience among leaders;

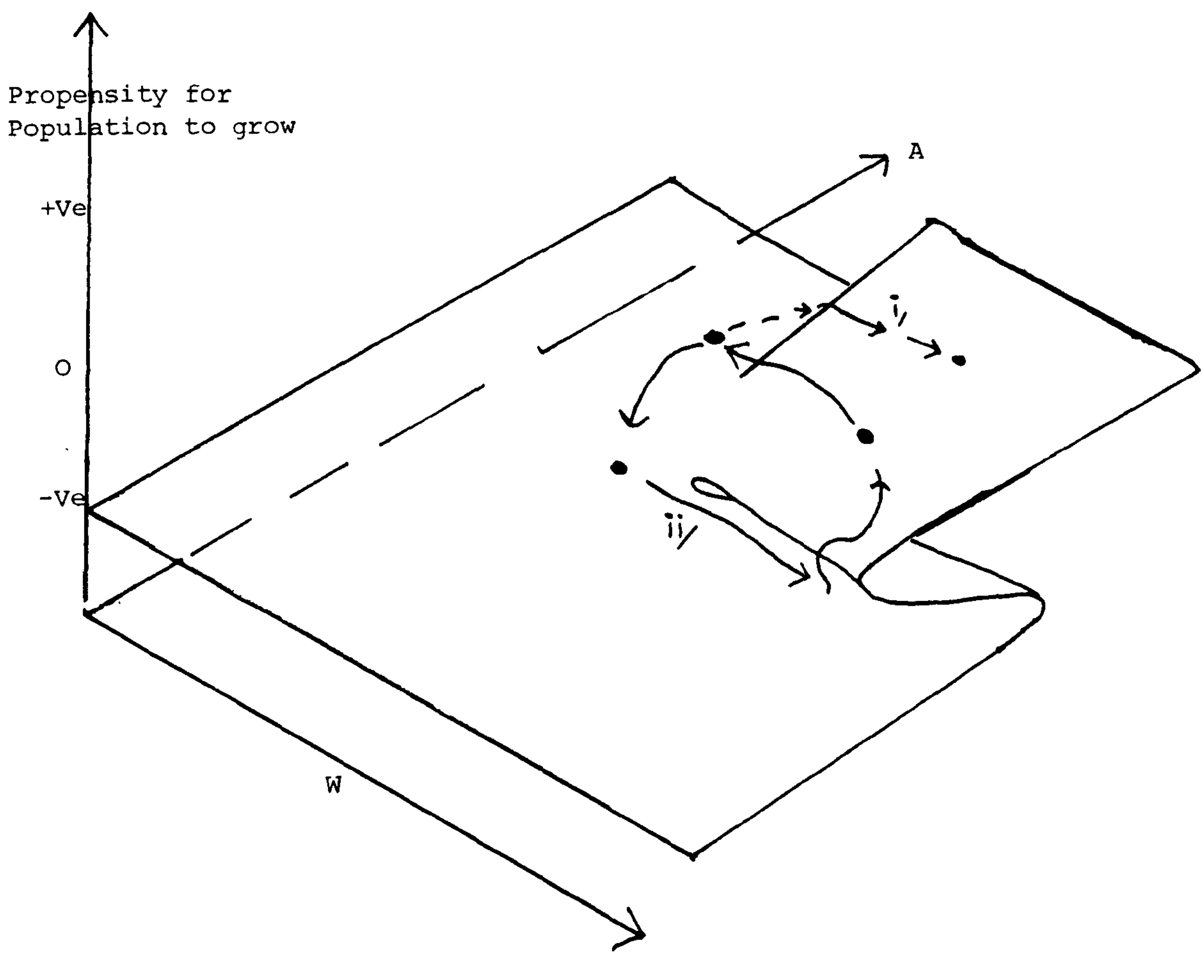
(iv) lack of appropriate attitudes amongst the populations to adjust to industrial non-familial based work;

(v) great poverty;

(vi) mainly rural (sometimes nomadic) population.

These make investment in the Third World countries by the developed nations an unattractive proposition. However, other aspects of ability have improved, in many cases by Western intervention, such as improved life expectancy, improved health and literacy (i.e. the physical quality of life index has risen (Hicks and Streeten, 1979)). For example, Gwatkin and Brandel state that there has been a fifty per cent increase in life

Figure 6.7.2.1 - Future Possibilities for Population Growth in the C.P.R.



The numbered paths correspond to the description in the text expectancy since the Second World War from less than 40 to over 55 years (although there is considerable variation amongst continents as Table 6.7.3.1 illustrates):

	1950	mid 1970s
Latin America	52	62
Asia	43	57
Africa	37	47

Table 6.7.3.1:
Expectation of life (in years) for developing nations

This improvement in life expectancy has occurred at a fast pace and Gwatkin and Brandel point out that what had taken Western Europe around seventy years to achieve in the nineteenth cetury has been accomplished in about fifteen years in the developing nations.

However, this improved mortality is acting to hamper other aspects of ability, particularly as it adds to the burden on food supply. Overall, the ability of the population is limited and this will limit the potential for sustained population growth.

Considering the desire for children in such societies, it is likely that this willingness will decrease with increasing Western influences disrupting wealth flows within the family (Caldwell, 1982) chiefly as a result of the provision of education. However, it is unlikely that willingness to have children will fall dramatically, as familial production is likely to persist for some considerable time as the dominant form of production, and the population is of a rural nature with little effective community support for old age or for insurance against risk.

Thus it would seem that the completion of a demographic transition will not occur for a considerable lead time (at least fifty years), although there are perhaps signs of the start of such a process of transition according to Dyson and Murphy (1985).

Also there is some urbanisation which is mainly centred around major trade ports, political capitals and Western owned primary extraction industries. (For an account of such urbanisation, see Boserup, 1981).

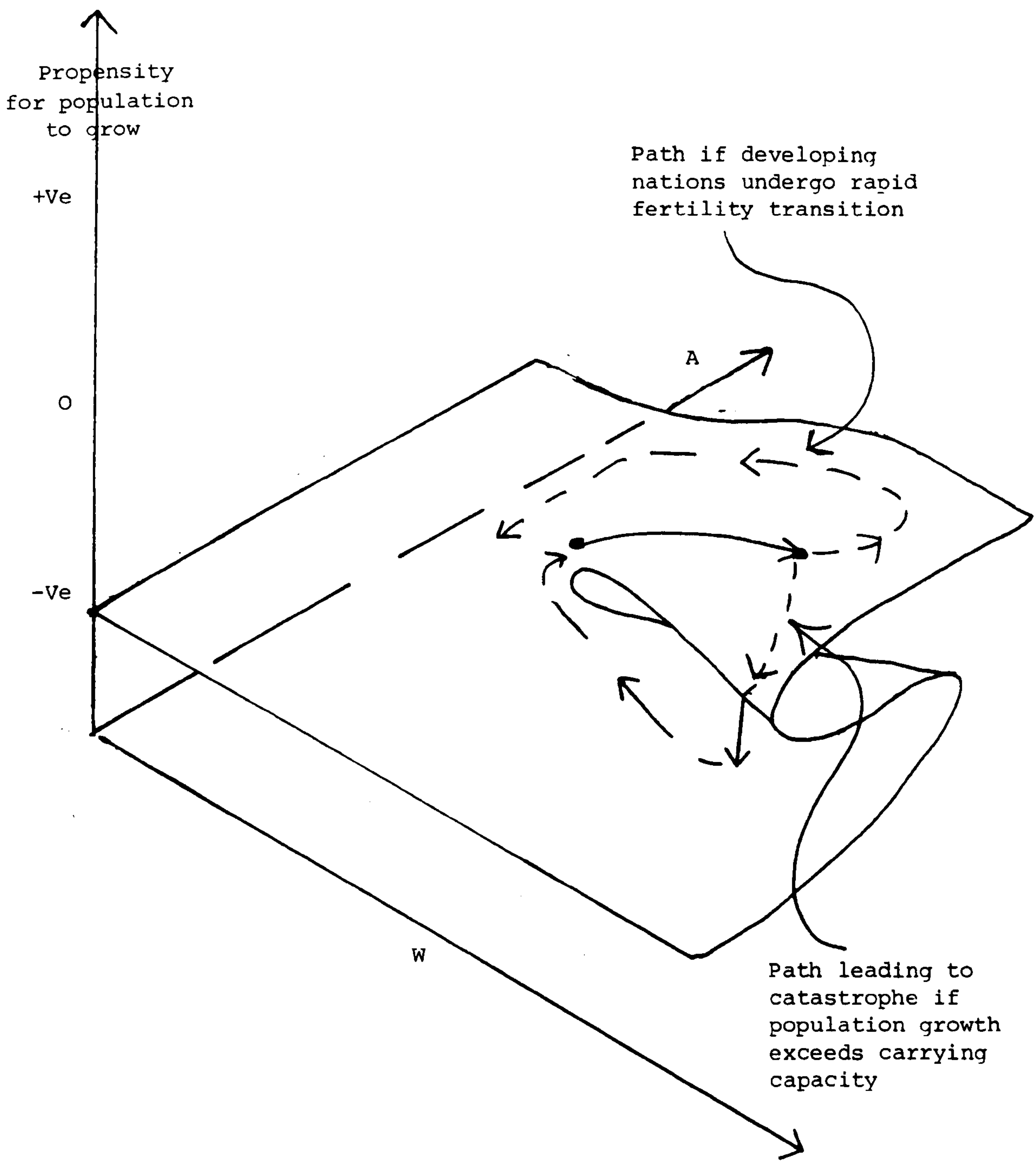
This discussion, based upon the catastrophe model, suggests that fast growth of population in the developing nations is likely to persist and growth will eventually be checked by Malthusian positive forces or exogenous mortality, which will result in a catastrophic fall in population size as ability cannot be increased fast enough. From such a position, renewed fast growth could occur then another catastrophic fall and so on leading to a saw-toothed pattern of boom and slump in the population. It would only be by allowing further Western influence, increasing investment and improving education, health and community care that willingness to have children could be reduced and catastrophe avoided.

If the influences on willingness to have children are insufficient to reduce fertility, then a simple geometric model with a consideration of an upper biological level of carrying capacity should be used for forecasting over the next fifty years.

The catastrophe map in Figure 6.7.3.1 illustrates the two possible scenarios, the one discussed above, and an optimistic scenario whee a point of inflection is encountered and a sigmoid is a suitable model. This scenario of a slowing of population growth is given some weight by the findings of Dyson and Murphy (1985) who refer to recent rises in fertility in these nations as a result of dietary improvements and medicines are giving rise to a "pre-fertility transition rise of fertility".

The division between the two scenarios and the effect of a

Figure 6.7.3.1 - Catastrophe Surface for the Future of Population Growth in Under Developed Nations



restricted ability will be made even more striking if a relative ability dimension is used. (Note that, if the willingness to have children cannot be altered, then the catastrophe map may be reduced to the simple fold as in Figure 6.4.2 with carrying capacity as the controlling factor.)

If the non-catastrophic path is followed, then ability must improve very quickly which seems unlikely, for even with, for instance, major wealth transfers from developed nations, as suggested by the Brandt Report (1980), and major fertility control programmes as advocated by Kirk (1984), it is unlikely that social and cultural systems could evolve quickly enough to cope with the economic changes.

The path towards catastrophe may have an even higher propensity if the drought and famines which struck Ethiopia and Sudan are evidence of climatic change which will render the already poor agricultural lands even more inhospitable. Hence ability is reduced and the population follows the path that encounters the cusp from above leading to a catastrophic fall.

Thus the methodology of Catastrophe Theory is useful in allowing one to make a subjective assessment of the quality of forecasts made by the suitable trend models. For advanced and emerging populations no reason can be found to refute the use of sigmoid models. However, there is greater uncertainty regarding underdeveloped populations, and the simpler exponential models are advised.

6.8 Conclusion

In this Chapter a method for holistically and parsimoniously modelling the propensity for population growth has been outlined. The methodology is developed by the use of catastrophe theory which, by using a qualitative approach, overcomes inherent problems which would arise if quantitative methods were employed. The methodology based upon catastrophe theory, it has been shown, can model successfully historical population growth and has allowed possible future scenarios for population in different regions of the world.

Hence one can model population change in the global type of context that Cowgill (1975) suggests should be used when he states:

> "Rather than seeing population growth as an inherent tendency of human population which is permitted by technological innovations, I see growth as a human possibility which is encouraged by certain institutional, as well as technological, or environmental circumstances. In general, asking whether population growth is the independent or the dependent variable is an inept question, and we should think of population variables as members of a set of variables, including technological and environmental variables and political, economic and other institutions which are all concomitantly interacting with one another".

By analysing likely movements of populations across the catastrophe surface, it becomes apparent that, for practical forecasting periods (up to fifty years), forecasts made by suitably constructed sigmoid models will give forecasts that may well give population sizes that are close to future observed population sizes. Certainly the discussions in this Chapter suggest that, excepting the underdeveloped populations, there is no reason to reject the use of sigmoid models and, providing adjustments are made for when inflections or escalations are encountered, then, at least for total population size, there is no reason to believe that forecasting performance of suitable sigmoids will be inferior to other forecasting methods (especially if a "cost" penalty is given to complexity and difficulty of construction).

The framework developed, apart from providing a useful heuristic tool by which the suitability of sigmoids and the likelihood of phase changes can be assessed, also allows assumptions to be identified and concisely displayed.

Much work remains to be done on the speculative model formed in this Chapter. Areas which require particular attention are in testing the model and to prescribe the limits of its applicability. Nevertheless the model developed is useful in adding to knowledge on population change. From the viewpoint of this thesis, the most important contributions that a Catastrophe Theory model gives to demography are:

(i) the ability to incorporate many influencing factors, and
(ii) to explain sudden change in patterns of population growth.

Chapter 7

Conclusions

In this chapter a summary of the main findings of this thesis are presented and demographic implications discussed. The aims of the thesis were established in the introduction, these were to reassess the suitability of using sigmoids to model and forecast human populations. This required an empirical investigation and a consideration of the rationality by using sigmoids. This reconsideration is in part stimulated by the increasing complexity and poor predictive performance of current demographic analysis.

In Chapter 2 assessment is made of the empirical and statistical suitability of sigmoids and other simple trend models for modelling and interpolation. Although it is recognised that the series of census counts of population are serially correlated, satisfactory models can be obtained by neglecting this feature and using a simple non-linear least squares routine based on minimisation of the proportionately weighted sum of squares.

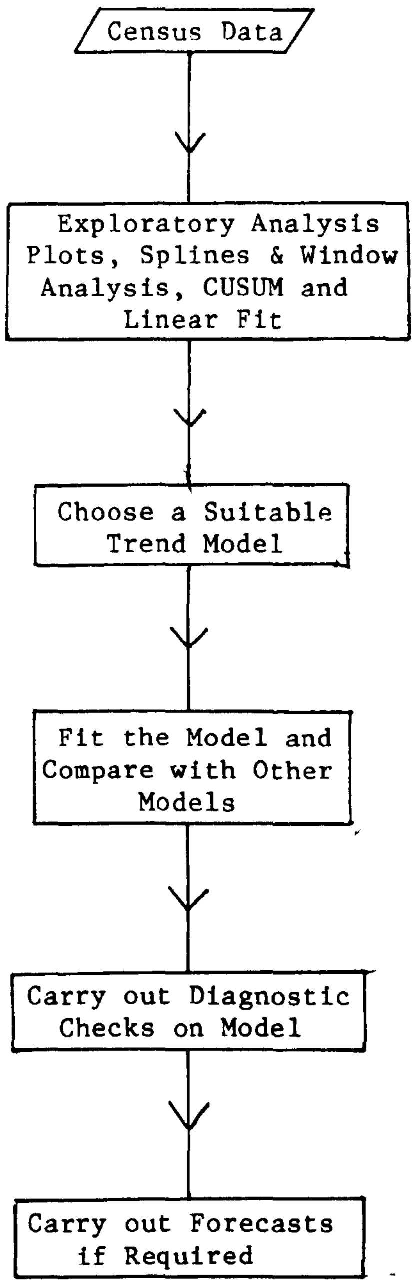
The actual fitting procedure developed is illustrated in the flowchart in Figure 7.1. In the exploratory analysis the Spline Analysis is a major contribution to assessing the suitability of sigmoid models and to give a bench mark by which the quality of fit can be judged.

Of the many parameterisations of the sigmoid models that were examined it was found that a logistic model of the form:-

$$P_t = \frac{P_\infty}{1+\left(\frac{P_\infty}{Po} - 1\right)} e^{-ct}$$

behaved closest to a linear model and so is considered the most suitable model. However, it did have significant parameter effects non linearities which precluded examination of the models by using tests developed for the linear model.

Figure 7.1 - The Fitting Procedure



The results from fitting trend models to a variety of populations in different social and cultural settings, indicate that the sigmoid models are for most populations the most suitable models, using criteria of closeness of fit, simplicity, demographic interpolation, parsimony and retrospective forecasting ability. In cases where sigmoids were not the most suitable models, exponentials often were. It is argued that this does not refute the case of sigmoids for such populations as it is merely a reflection of the artefact that the least squares cannot differentiate between sigmoid and exponential models until the population is significantly mature.

Difficulties did arise in several populations where there was a radical change in growth - labelled in the thesis as a phase change. This has the effect of rendering the above modelling approach inadequate for such populations.

To tackle these situations it was found best to model before and after this "phase change" with separate models. Detection of the need for such discontinuous modelling is made when exploring the data, but this is "after the event" detection and so is unable to aid forecasting decision making.

The conclusion after empirically modelling many populations, is that sigmoids, particularly logistic models, are good models of historical population time series and that populations for which sigmoid or exponential models are not suitable should be seen as exceptional.

Given that in developing nations, modelling is likely to be limited by the quality and quantity of data, sigmoid models have great potential to aid demographic analysis.

In Chapter 3 forecasting is examined and it is found that forecasts generated by extrapolation from the time series parameterised models are not very satisfactory as forecasts were often bettered by even simpler models. Other time series models, such as ARIMA models, are also found to be unsatisfactory giving poor forecasts and many being statistically questionable. A better approach is to recast the models in a recursive parametric form and to update parameters and generate forecasts using Kalman filtering. Doing this allowed confidence intervals on forecasts to be approximately set and retrospective forecasts to be made for the USA and the European populations examined. These forecasts are more accurate than other forecasts examined.

Some of the future forecasts compare closely with official projections, others differ to an appreciable degree, as is the case for the USA, Sweden and Scotland. Which are the most likely can only be answered in time.

The conclusion of this chapter is that sigmoid models (the logistics were the only ones examined) are good, when used in a Kalman filter procedure, for generating forecasts of total population. It is also found that providing there are no phase changes then satisfactory forecasts for lead times by up to fifty years can be obtained.

Attention is turned to modelling and forecasting components in Chapter 4. It is found that sigmoid models can be used in conjunction with ARMA time series models in a "top down" approach to produce satisfactory models and forecasts of the number of births per year. These forecasts can be equipped with confidence intervals. These were generated for several of the developed nations and using these forecasts the early detection of the onset of a "phase change" is apparent when there is deviation from expected paths when future observations are recorded. A useful 'by product' of this analysis is the finding that improvements in life expectancy over time can be modelled by sigmoid models.

Moving average models gave good expressions and sensible forecasts of the crude death rate series and hence an indication of migration levels can be obtained from

$$M_t = P_t - P_{t-1} - B_t + D_t$$

Logistic based forecasts of births that were generated from the total population can be combined with the probability of surviving to the next age group and the numbers in each age group to produce age specific (and sex specific) forecasts of the population. The amalgamation of these forecasts compared favourably to the forecasts made in Chapter 3 and retrospective forecasts compared well with observed population counts. These models also acted to support and verify the models of total population. Hence two modelling and forecasting methodologies are incorporated into one framework which has the advantages of simplicity and produces long term forecasts which sum to anticipated levels.

To sum up, in Chapters 2, 3 and 4 it is demonstrated empirically that for many populations sigmoids are useful for:-

(a) modelling and interpolating human populations at the national level;

(b) forecasting national populations up to a lead time of fifty years;

(c) modelling and forecasting component populations and to underpin bottom up component models;

(d) giving confidence limits to forecasts.

Hence, on empirical grounds a reconsideration of sigmoid models is worthwhile and will improve interpolation and prediction in demography. In Chapter 5 the possibility of linking the use of sigmoids with demographic theory was investigated and the rationality of using sigmoids was considered.

The notion of sigmoids as a law of population growth is quickly dismissed. Yet the suitability of sigmoid models to reflect population development for such a wide diversity of cultures and settings must be considered to be more than simply a fortunate coincidence of events.

It was shown that sigmoid models can be viewed as a mathematical expression of Malthusian theory in which the preventative check is strong. (Exponentials being more likely to be suitable when the positive check dominates). Hence, if Malthusian based theory is appropriate then it is rational to use sigmoid models.

Explanations of regulatory mechanisms based on population pressure and carrying capacity are found to be unsuitable but after a historical review of population development, evidence emerges of Malthusian forces, although these have been socially and culturally mediated. Thus one conclusion is that the use of sigmoid models are theoretically sound if viewed in a Malthusian framework. However, it was also pointed out that little is to be gained from attaching theoretical interpretations to the model parameters.

A feature which emerged from the discussion in Chapter 5 is that, as populations have evolved, explanations of the regulating mechanisms have changed. To understand how different regulating mechanisms combine to effect populations, the system in which the population operates must be viewed holistically. This was investigated in Chapter 6 and it is demonstrated that conventional demographic modelling which focusses on subsets of the factors effecting population is inadequate. A speculative inquiry into the appropriateness of using Catastrophe Theory gives encouraging results when analysing the processes involved in demographic change. The analysis showed how abrupt "phase changes" in growth can come about through slowly changing variables. For this reason and for its ability to deal holistically and qualitatively with population change the use of Catastrophe Theory as a qualitative model was developed and found to be useful in portraying the development of several historical populations, notably Ireland and France which have posed problematic quantitative modelling. The model developed is found to be useful in allowing one to assess the suitability of using sigmoid models and to speculate on the likelihood of "phase changes". The catastrophe framework also allows one to pictorically state assumptions. This model allowed speculation to be made in general terms of future scenarios for population in different regions of the world and gives no evidence to refute the use of sigmoids.

In short the findings of this thesis are that sigmoid models are empirically suitable as interpolative models of population and as a base for forecasting and that the models are on sound theoretical ground.

The advantages available to population modelling and forecasting of these findings are simplicity, speed and improved accuracy of interpolation and forecasts. Top down and bottom up approaches can be integrated to produce more self consistent forecasts and when viewed in a Catastrophe Theory frame the models and forecasts can be assessed in a wider context. It is an important gain to be able to equip forecasts with confidence intervals, these are wide and highlight the uncertainty which future planners face.

It has also been illustrated in this thesis that many of the criticisms levelled against sigmoids as models were either unfounded or unjust. So by ignoring sigmoid models modern demographic analysis neglects a useful tool. Therefore it is worthwhile for demographers to reconsider the adoption of "top down" approach as has been expedited in this thesis.

A forecasting scheme based on such a simplified approach would allow more attention to be devoted to a discussion of possible scenarios and the true inherent uncertainty to be emphasised. (The lack of such features are viewed by Keilman and Cruijsen (1984) as deficiences with contemporary modelling and forecasting schemes). Catastrophe Theory models would further help in this analysis.

To conclude, it is advocated in this thesis that in conjunction with conventional "bottom up" demographic based models that sigmoid models of the historical population should be obtained. From these models coupled with inputs from bottom up models, models of vital population statistics can be derived and age-sex specific models and forecasts obtained. By employing these means (all be it via Kalman filters) more rationally sound forecasts can be made which have with them a measure of uncertainty. Forecasts are made under the assumption of no sudden changes in growth patterns. If a phase change occurs then re-modelling is required. These disturbances can be detected after occurance by the exploratory analysis described in Chapter 2, by comparing future vital events to forecasts and by monitoring the residual variance when Kalman filtering is applied. Before the event detection is best carried out in a qualitative framework using a Catastrophe Theory model to consider how gradual changes occurring in the population system might manifest themselves in population growth rates.

BIBLIOGRAPHY

Adby, P.R. and Dempster, M.A.H. (1974), "Introduction to Optimisation Methods", Chapman and Hall, London.

Ahlburg, D.A. (1982), "How Accurate are the U.S. Bureau of the Census Projections of Total Live Births?", J. Forecasting, 1, 365-374.

Allen, Sir Geoffrey (1985), "Translation of Technological Developments into Commercial Products or Processes" - "Technology, business and society", Lecture at City University, Table from Futurescope, Section I, July 1983, 1-21, Arthur D. Little Decision Resources.

Armstrong, P.J. (1985), "Work, Rest or Play? Changes in Time Spent at Work", Technical Change Centre, London, 26-44.

Ascher, W. (1979), "Forecasting: An Appraisal for Policy Makers and Planners", The John S. Hopkins University Press, Baltimore and London.

Austin A.L. and Brewer, J.W. (1971), "World Population Growth and Related Technical Problems", Technology Forecasting and Social Change, 3, 23-49.

Barney, G.O. (ed.) (1980), "The Global 2000 Report to the President of the U.S.", Elmsford, N.Y., Pergamon.

Baron, Y.M. (1980), "The Case for Differences in the Responses of Humans and other Animals to High Density", in Cohen et al. (1980), 247-273.

Bartlett, M.S. (1970), "Stochastic Population Models in Ecology and Epidemiology", Methuen and Co. Ltd., London.

Bates, D.M. and Watts, D.G. (1980), "Relative Curvature Measures of Nonlinearity", J.R. Statist. Soc., Ser. B., 42, 1-25.

Becker, G.S. (1965), "A Theory of the Allocation of Time", Econ.

J., 75, 493-517.

Becker, G.S. (1974), "A Theory of Social Interactions", J. Political Economy, 82, 1063-1093.

Beckman, L. (1978), "Couples' Decision-making Processes regarding Fertility", in "Studies in Population" by Taeuber K. et al., Academic Press, London, 57-81.

de Beer, J. (1983), "A Time Series Model for Cohort Data", Netherlands Central Bur. of Statist., BPA No. 10434-83-ml.

Bell, D. (1974), "The Coming of Post-industrial Society", Heinemann, London.

Bell, D. (1979), "The Cultural Contradictions of Capitalism", Heinemann, London.

Bhattacharya, C.G. (1966), "Fitting a Class of Growth Curves", Sankhya, Ser. B, 28, parts 1 & 2, 1-10.

Birdsall, N. (1983), "Fertility and Economic Change in Nineteenth Century Europe. A Comment", Pop. and Development Rev., 9, 111-123.

Birdsell, J.B. (1968), "Some Predictions for the Pleistocene based Equilibrium Systems among Recent Hunter-gatherers", in "Man the Hunter", eds. R.B. Lee and I. de Vore, Aldine, Chicago, 229-240.

Birdsell, J.B. (1978), "Spacing Mechanisms and Adaptive Behaviour of Australian Aborigines", in "Population Control by Social Behaviour", eds. F.J. Ebling and D.M. Stoddart, Inst. of Biol., London, 213-244.

Bonar, J. (1966), "Malthus and his Work", Frank Cass, London.

Boserup, E. (1981), "Population and Technology", Blackwell, Oxford.

Boserup, E. (1986), "Shifts in the Determinants of Fertility in the Developing World: Environmental, Technical, Economic and Cultural Factors", in "The State of Population Theory", ed. D. Coleman and R. Schofield.

Boulding, K.E. (1973), "Zoom, Gloom, Doom and Room", The New Republic, 169, 25-27.

Bourgeois-Pichart, J. (1981), "Recent Demographic Change in Western Europe: An Assessment", Pop. and Development Rev., 7, 19-42.

Bowley, A.L. (1924), "Births and Population of Great Britain", J.R. Econ. Soc., 34, 188-192.

Box, G.E.P. and Cox, D.R. (1964), "An Analysis of Transformations", J.R. Statist. Soc., Ser. B, 26, 211-252.

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

Box, M.J. (1971), "Bias in Nonlinear Estimation", J.R. Statist. Soc., Ser. B, 33, 171-201.

Boyce, W.E. and di Prima, R.C. (1977), "Elementary Differential Equations", 3rd edn., Wiley, New York.

Brand, R.J., Pinnock, D.E. and Jackson, K.L. (1973), "Large Sample Confidence Bounds for the Logistic Response Curve and its Inverse", Amer. Statistician, 27 157-160.

Brandt, W. (1980), "North-South: A Programme for Survival", Pan Books, London.

Brass, W. (1974), "Perspectives in Population Prediction Illustrated by the Statistics of England and Wales", J.R. Statist. Soc., Ser. A, 137, 532.

Brass, W. (1979), "Population Projections for Planning and Policy", C.P.S. Working Paper No. 79-1, Centre for

Population Studies, London School of Hygiene and Tropical Medicine, London University.

Braverman, H. (1974), "Labor and Monopoly Capital", Monthly Rev. Press, New York.

Broecker, W.S. et al., (1985), "Does the Ocean-atmosphere System have more than one Stable Mode of Operation?", Nature, 315, 21-26.

Buckland, S.T. (1985), "Calculation of Monte Carlo Confidence Intervals, Algorithm AS 214", Appl. Statist. 34, 296-301.

Bulatao, R.A. and Lee, R.D. (1983), "Determinants of Fertility in Developing Countries", Vol. I and II, Academic Press, New York.

Bulmer, M.G. (1975), "The Statistical Analysis of Density Dependence", Biometrics, 31, 901-911.

Burch, E.S. (1972), "The Caribou/Wild Reindeer as a Human Resource", Amer. Antiquity, 37, 339-368.

Burghes, D.N. and Borrie, M.S. (1981), "Modelling with Differential Equations", published by Ellis Horwood, distributed by John Wiley & Sons, Chichester, 14-21, 63, 64.

Butz, W.P. and Ward, M.P. (1979), "The Emergence of Countercyclical U.S. Fertility", Amer. Econ. Rev., 69, 318-328.

Butz, W. and Ward, M. (1979), "Will U.S. Fertility Remain Low?", Pop. and Development Rev., 5, 663-688.

Butzer, K.W. (1981), "Comment on Hayden's 1981 Paper", Current Anthropology, 22, 531.

Cain, M. (1981), "Risk and Insurance: Perspectives on Fertility and Agrarian Change in India and Bangladesh", Pop. and

Development Rev., 7, 435-474.

Cain, M. (1982), "Perspectives on Family and Fertility in Developing Countries", Population Studies, 36, 159-175.

Cain, M. (1983), "Fertility as an Adjustment to Risk", Pop. and Development Rev. 9, 688-702.

Cain, M. (1985), "On the Relationship between Landholding and Fertility", Population Studies, 39, 5-15.

Cain, M. (1986), "Risk and Fertility: A Reply to Robinson", Population Studies, 40, 299-304.

Caldwell, J.C. (1982), "Theory of Fertility Decline", Academic Press, London.

Calot, G. (1981), "The Demographic Situation in France", Pop. Trends, 25, 15-20.

Calot, G. and Blayo, C. (1982), "Recent Course of Fertility in Western Europe", Population Studies, 36, 349-372.

Cameron, R. (1973), "The Logistics of European Economic Growth. A Note on a Historical Periodisation", J. European Econ. History, 2, 145-148.

Cannan, E. (1895), "The Probability of a Cessation of the Growth of Population in England and Wales during the Next Century", Econ. J., 5, 505-515.

Carneiro, R.L. (1970), "A Theory of the Origin of the State", Science, 169, 733-738.

Carr-Saunders, A. (1922), "The Population Problem: A Study in Human Evolution", Clarendon Press, London.

Carter, L.R. and Lee, R.D. (1981), "Joint Forecasts of U.S. Marital Fertility, Nuptiality, Births and Marriages using Time Series Models", J. Amer. Statist. Assoc., 81, 902-911.

C.P.R.S. (Central Policy Review Staff) (1977), "Population and the Social Services", H.M.S.O., London.

Central Statistical Office (1959 and 1982), "Annual Abstract of Statistics", H.M.S.O., London.

Chen Kuan, I. (1982), "China's Population Policy", Current History, 251-254, 275-277.

Clark, C. (1977), "Population Growth and Land Use", Macmillan Press Ltd., London.

Clarke, J.I. (1973), "Pressure on Resources: The Problem of Evaluation", in "Resources and Population", eds. B. Benjamin, P.R. Cox and J. Peel, Academic Press, London.

Coale, A.J. (1974), "The History of the Human Population", Scientific Amer., 231, 41-51.

Coale, A.J. (1985), "Nuclear War and Demographers' Projections", Pop. and Development Rev., 11, 483-493.

Cohen, M.N. et al. (eds.) (1980), "Biosocial Mechanisms of Population Regulation", Yale University Press, London.

Cohen, M.N. (1980), "Speculations on the Evolution of Density Measurement and Population Regulation in Homo Sapiens", in Cohen et al. (1980), 275-303.

Coleman, D.A. (1980), "Recent Trends in Marriage and Divorce in Britain and Europe", in Hiorns, 1980, 83-124.

Coleman, D. (1986), "Population Regulation - a Long-range View", in "The State of Population Theory", eds. D. Coleman and R. Schofield, Blackwell, Oxford.

Coleman, D. and Schofield, R. (eds.) (1986), "The State of Population Theory", Blackwell, Oxford.

Cowgill, D.O. (1949), "The Theory of Population Growth Cycles", Amer. J. Sociology, 55, 163-70.

Cowgill, G.L. (1969), "Population Pressure - As a Non-explanation", in "Population Evolution and Birth Control", eds. G. Hardin and G.L. Cowgill, W.H. Freeman, San Francisco.

Cowgill, G. (1975), "On Causes and Consequences of Ancient and Modern Population Changes", Amer. Anthropologist, 77, 505-525.

Cox, D.R. et al. (1983), "Time Series Analysis: Short Course Notes", Dept. of Mathematics, Imperial College, London.

Cox, P.R. (1970), "Demography", 4th Edition, Cambridge University Press.

Daly, H.E. (1985), "Marx and Malthus in North-east Brazil: A Note on the World's Largest Class Difference in Fertility and its Recent Trends", Population Studies, 39, 329-338.

Deevey, Jr., E.S. (1960), "The Human Population", Scientific Amer., 203, 195-204.

Diggle, P.J. (1985), "Comment on Silverman's Paper", J.R. Statist. Soc., Ser. A, 47, Part 1, 28.

Dorn, H.F. (1950), "Pitfalls in Population Forecasts and Projections", J. Amer. Statist. Assoc. 45, 311-334.

Douglas, M. (1966), "Population Control in Primitive Groups", Brit. J. Sociology, 17, 263-273.

Dryer, C. (1978), "Population and Society in Twentieth Century France", Hodder and Stoughton, London.

Dyson, T. (1977), "The Demography of the Hadza in Historical Perspective", Proc. African Historical Demography Conf., Centre for African Studies, Edinburgh.

Dyson, T. and Murphy, M. (1985), "The Onset of Fertility Transition", Pop. and Development Rev., 11, 399-440.

E.C.S.L./C.A.P.S. (1980), The Lucas Institute for Engineering Production, Univ. of Birmingham.

Easterlin, R.A. (1968), "Population, Labor Force and Long Swings in Economic Growth", National Bureau of Econ. Res., Columbia University Press, New York.

Easterlin, R.A. (1976), "The Conflict between Aspiration and Resources", Pop. and Development Rev., 2, 417-425.

Easterlin, R.A. (1978), "What will 1984 be like? Socioeconomic Implications of Recent Twists in Age Structure", Demography, 15, 397-432.

Easterlin, R.A. (1980), "Birth and Fortune", Grant McIntyre, London.

Ehrenberg, A.S.C. (1982), "How Good Is Best?", J.R. Statist. Soc., Ser. A, 145, 364-366.

Elandt-Johnson, R.C. and Johnson, N.L. (1980), "Survival Models and Data Analysis", Wiley, New York.

Epstein, Y.M. (1980), "Physiological Effects of Crowding on Humans", in Cohen et al., 209-224.

Ermisch, J. (1981), "An Emerging Secular Rise in the Western World's Fertility?", Pop. and Development Rev., 7, 677-684.

Ermisch, J. (1982), "Economic Models of Period Fertility in Britain", Policy Studies Institute, London.

Ermisch, J.F. (1983), "The Political Economy of Demographic Change", Heinemann, London.

Ermisch, J. (1984), "Long Term Forecasting of Housing Demand in Great Britain: The Importance of Demographic Factors", presented at 4th Int. Symposium on Forecasting, London.

Ermisch, J. and Joshi, H. (1983), "Modelling Fertility Patterns over Time: A Progress Report", Policy Studies Institute, London.

Feller, W. (1940), "On the Logistic Law of Growth and its Empirical Verifications in Biology", Acta Biotheoretica, Ser. A, 51-66.

Feller, W. (1950), "An Introduction to Proability Theory and its Applications", Vol. 1, Wiley, Chichester.

Flew, A. (ed.), (1982), "Thomas Malthus: An Essay on the Principle of Population", Penguin, Middlesex.

Freedman, D.S. (1963), "The Relation of Economic Status to Fertility", Amer. Econ. Rev., 53, 414-426.

Freedman, J.L. (1980), "Human Reactions to Population Density", in Cohen et al., 189-207.

Freeman, C. (1981), "Technical Innovation and Long Waves in World Economic Development", Futures - Special Issue, 13.

Freeman, C. (1983), "Long Waves in the World Economy", Butterworth, London.

Freeman, C. et al. (1982), "Unemployment and Technical Innovation", Frances Pinter, London.

Fujita, W.M. and Utida, S. (1953), "The Effect of Population Density on the Growth of Animal Populations", Ecology, 34, 488-498.

Galle, O. and Gove, W. (1978), "Overcrowding, Isolation and Human Behaviour: Exploring the Extremes in Population Distribution", in "Studies in Population" by K. Taeuber et al., 95-132, Academic Press, London.

Galloway, P.R. (1985), "Annual Variations in Deaths by Cause, Prices and Weather in London 1670-1830", Population Studies, 39, 487-505.

Galloway, P.R. (1986), "Long Term Fluctuations in Climate and Population in the Pre-industrial Era:, Pop. and Development Rev., 12, 1-24.

GENSTAT 4.01 (1977), Rothamsted Experimental Station.

Gershuny, J. (1978), "After Industrial Society", Macmillan Press Ltd., London.

Gill, D.E. (1972), "Density Dependence and Population Regulation in Laboratory Cultures of Paramecium", Ecology, 53, 701-708.

Gilland, B. (1983), "Considerations on World Population and Food Supply", Pop. and Development Rev., 9, 203-211.

Gilmore, R. (1981), "Catastrophe Theory for Scientists and Engineers", Wiley, New York.

Girard, A. and Roussel L. (1979), "Fecondite et Conjoncture", Population 5, 567-588.

Glasbey, C.A. (1979), "Correlated Residuals in Non-linear Regression applied to Growth Data", Appl. Statist., 28, 251-259.

Glasbey, C.A. (1980), "Nonlinear Regression with Autoregressive Time Series Errors", Biometrics 36, 135-140.

Glass, D.V. (ed.) (1953), "Introduction to Malthus", Watts & Co., London.

Glass, D.V. (1967), "Demographic Prediction", The Third Royal Society Nuffield Lecture, Proc. Roy. Soc., Ser. B, 168, 119-139.

Goldstone, J.A. (1986), "The Demographic Revolution in England: a Re-examination", Population Studies 49, 5-33.

Gompertz, B. (1825), "On the Nature of the Function Expressive of the Law of Human Mortality and on a New Method of

Determining the Value of Life Contingencies", Phil. Trans. Roy. Soc., 513-585.

Goodyear, A.C. (1981), "Comment on Hayden's 1981 Paper", Current Anthropology, 22, 533.

Greene, W.H. and Quester, A.O. (1982), "Divorce Risk and Wives' Labour Supply", Social Science Quarterly, 63, 16-27.

Grigg, D.B. (1980), "Population Growth and Agrarian Change: An historical perspective", Cambridge University Press.

Guckenheimer, J. (1978), "The Catastrophe Controversy", The Mathematical Intelligencer, 1, 15-20.

Gumerman, G.J. (1975), "Alternative Cultural Models of Demographic Change: Southwestern Examples", in "Population Studies in Archaeology and Biological Anthropology: A Symposium", ed. A.C. Swedlund, Amer. Antiquity, 40, 104-115.

Gutman, M.P. and Leboutte, R. (1984), "Rethinking Protoindustrialisation and the Family", J. Interdisciplinary History, 24, 587-608.

Gwatkin, D.R. and Brandel, S.K. (1982), "Life Expectancy and Population Growth in the Third World", Scientific Amer., 246.

Habakkuk, H.J. (1972), "Population Growth and Economic Development since 1750", Leicester University Press.

Hajnal, J. (1955), "The Prospects for Population Forecasts", J. Amer. Statist. Assoc., 50, 309-322.

Hajnal, J. (1982), "Two Kinds of Pre-industrial Household Formation System", Pop. and Development Rev. 8, 449-494.

Harrison, P.J. and Stevens, C.F. (1976), "Bayesian Forecasting" (with discussion), J. R. Statist. Soc., Ser. B, 38, 205-247.

Hart, H. and Hertz, H. (1944), "Expectation of Life as an Index of Social Progress", Amer. Sociological Rev., 9, 609-621.

Hart, H. (1945), "Logistic Social Trends", Amer. J. Sociology, 50. 337-352.

Hassan, F. (1973), "On Mechanisms of Population Growth during the Neolithic", Current Anthropology, 14, 535-542.

Hassan, F.A. (1980), "The Growth and Regulation of Human Population in Prehistoric Times", in Cohen et al., 305-319.

Hassan, F.A. (1981), "Comment on Hayden's 1981 Paper", Current Anthropology, 22, 534.

Hatcher, J. (1977), "Plague Population and the English Economy 1348-1530", Macmillan, London.

Hayden, B. (1975), "The Carrying Capacity Dilemma", in "Population Studies in Archaeology and Biological Anthropology: A Symposium", ed. A.C. Swedlund, Amer. Antiquity, 40, 11-21.

Hayden, B. (1981), "Research and Development in the Stone Age: Technological Transitions among Hunter Gatherers", Current Anthropology, 22, 519-548.

Hicks, N. and Streeten, P (1979), "Indicators of Development - the Search for a Basic Yardstick", World Development, 7, 567-580.

Hiorns, R.W. (ed.) (1980), "Demographic Patterns in Developed Societies", Taylor & Francis, London.

Hollingsworth, T.H. (1973), "Population Crises in the Past", in "Resources and Population", eds. B. Benjamin, P.R. Cox and J. Peel, Academic Press, London.

Hotelling, H. (1927), "Differential Equation subject to Error and Population Estimates", J. Amer. Statist. Assoc., 27, 283-314.

Howell, N. (1986), "Feedbacks and Buffers in Relation to Scarcity and Abundance: Studies of Hunter-Gatherer Populations", in "The State of Population Theory", eds. D. Coleman and R. Schofield.

Huss, M. (1980), "Demographic Public Opinion and Politics in France 1974-80", Occasional paper No. 16, Dept. of Geography, Queen Mary College, University of London.

I.N.S.E.E. (1977), "Annuaire Statistique de la France - Resultats de 1975", Institut National de la Statistique et des Etudes Economiques, 82.

Jenkins, C. and Sherman, B. (1979), "The Collapse of Work", Eyre, Methuen, London.

Jensen, A.L. (1975) "Comparison of Logistic Equations for Population Growth", Biometrics 31, 853-862.

Jones, K.F. and Kay, N. (1979), "Pearl's Law and the Growth of the Scottish National Party", BIAS, 6. 195-214.

Joshi, H. and Overton, E. (1984), "The Female Labour Force in Britain 1971-1991", CPS Research paper No. 84-1.

Karlin, R.A. (1980), "Social Effects of Crowding on Humans", in Cohen et al., 225-245.

Keilman, N. and Cruijsen, H. (1984), "Population Forecasting - A Systems View", Fourth Int. Symposium on Forecasting, 8-11 July, London.

Keyfitz, N. (1968), "Introduction to the Mathematics of Population with Revisions", Harvard University, New York.

Keyfitz, N. (1972), "On Future Population", J. Amer. Statist. Assoc., 67, 347-363.

Keyfitz, N. (1979), "Prospects of Population: Methodology and

Assumptions", ST/ESA/SER.A/67, United Nations, New York.

Keyfitz, N. (1981), "The Limits of Population Forecasting", Pop. and Development Rev., 7, 379-593.

Keyfitz, N. (1982), "Population Change and Social Policy", Abt Books, Cambridge, Mass..

Keyfitz, N. and Flieger, W. (1968), "World Population - An Analysis of Vital Data", University of Chicago Press Ltd., Chicago.

Keyfitz, N. and Flieger, Q. (1971), "Population: Facts and Methods of Demography", W.H. Freeman & Co., San Francisco.

Kiernan, K.E. and Eldridge, S.M. (1985), "A Demographic Analysis of First Marriages in England and Wales: 1850-1980", Centre for Population Studies, Research paper 85-1, London School of Hygiene and Tropical Medicine, London.

Kirk, M. (1984), "The Return of Malthus?", Futures, 16, 124-138.

Lack, D. (1954), "The Natural Regulation of Animal Numbers", Clarendon Press, Oxford.

Lack, D. (1966), "Population Studies of Birds", Clarendon Press, Oxford.

Lamb, H.H. (1982), "Climate, History and the Modern World", Methuen, London.

Land, K.C. (1986), "Methods for National Population Forecasts: A Review", J. Amer. Statist. Soc., 81, 888-901.

Land, K.C. and Cantor, D. (1983), "AR & MA Models of Seasonal Variation in U.S. Birth and Death Rates", Demography, 20, 541-568.

Langer, W. (1964), "The Black Death", in "Man and the Ecosphere", ed. P.R. Ehrlich, J.P. Holdren and R.W. Hold (1971), W.H. Freeman & Co., San Francisco.

Larson, D.A. and Wilford, W.T. (1979), "The Physical Quality of Life. Index: A Useful Social Indicator?", World

Development, 7, 581-584.

Leach, D. (1981), "Re-evaluation of the Logistic Curve for Human Populations", J.R. Statist. A, 144, Part 1, 94-103.

Leach, D. (1985), "Production, Productivity and Employment: Models for U.K. Index of Production Industries, 1950-1990", J. Interdisciplinary Econ., 1, 29-42.

Leach, D. and Wagstaff, H. (1986), "Future Employment and Technological Change", Kogan Page, London.

Lee, R. (1973), "Population in Pre-industrial England: An Econometric Analysis", Quarterly J. Econ. 87, 581-607.

Lee R.B. (1980), "Lactation, Ovulation, Infanticide and Women's Work: A Study of Hunter-Gatherer Population Regulation", in Cohen et al.

Lee, R.D. (1974), "Forecasting Births in Post-Transition Populations: Stochastic Renewal with Serially Correlated Fertility", J. Amer. Statist. 69, 607-617.

Lee, R.D. (1974), "The Formal Dynamics of Controlled Populations. And the Echo, the Boom and the Bust", Demography, 11.

Lee R.D. (1979), "New Methods for Forecasting Fertility: An Overview", in "Prospects of Population: Methodology and Assumptions", Papers of the ad hoc group of experts in Demographic Projections, UN Population Studies No. 67, United Nations, New York, 215-221.

Lee, R.D. (1986), "Population Pressures and Responses Malthus & Boserup: A Dynamic Synthesis", in "The State of Population Theory", eds. D. Coleman and R. Schofield.

Leontief, W. (1983), "Technological Advance, Economic Growth and the Distribution of Income", Pop. and Development Rev. 9, 403-410.

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

Leslie, P.H. (1948), "Some Further Notes on the Use of Matrices in Population Mathematics", Biometrika, 35, 213-245.

Lesthaeghe, R. (1980), "On the Social Control of Human Reproduction", Pop. and Development Rev., 6, 527-548.

Lesthaeghe, R. (1983), "A Century of Demographic and Cultural Change in Western Europe: An Exploration of Underlying Dimensions", IPD Working Paper 1983-1, Vrije Universiteit, Brussels.

Lesthaeghe, R. (1986), "On the Adaptation of Sub-Saharan System of Reproduction", in "The State of Population Theory", eds. D. Coleman and R. Schofield.

Liebenstein, H. (1974), "Socio-economic Fertility Theories and their Relevance to Population Policy", Int. Labour Rev. 109, 443-457.

Lindert, P.H. (1983), "English Living Standards, Population Growth and Wrigley-Schofield", Explorations in Econ. History, 20, 131-155.

Lloyd, P.J. (1967), "American, German and British Antecedents to Pearl and Reed's Logistic Curve", Population Studies, 21, 99-108.

Long, J.F. and McMillen, D.B. (1984), "A Survey of Census Bureau Population Projection Methods", presented at the 4th Int. Symposium on Forecasting, London, July 1984, and the SSRC Conference on Forecasting in the Social and Natural Sciences, Boulder, Colorado, June 8-11,1984.

Lotka, A.J. (1925), "Elements of Physical Biology", Williams and Wilkins, Baltimore.

Lotka, A.J. (1931), "The Structure of a Growing Population", Human Biol., 3, 459-493.

Lotka, A.J. (1934), "On an Integral Equation in Population Analysis", Annals of Mathematical Statistics, 10, 144-161.

Mackintosh, J.H. (1978), "The Experimental Analysis of Overcrowding", in "Population Control by Social Behaviour", eds. F.J. Ebling and D.M. Stoddart, Inst. of Biology, London, 157-180.

McDonald, J. (1979), "A Time Series Approach to Forecasting Australian Total Live Births", Demography, 16, 575-601.

McDonald, J. (1980), "Births Time Series Models and Structural Interpretations", J. Amer. Statist. Assoc., 75, 39-41.

McEvedy, C. and Jones, R. (1978), "Atlas of World Population History", Penguin, Middlesex.

McKeown, T. (1976), "The Modern Rise of Population", Edward Arnold, London.

McLean, M.C. and Turner, A.N. (1937), "The Logistic applied to Canada's Population", Canadian J. Econ. and Political Science, 3, 241-248.

McNeil, D.R. (1974), "Pearl-Reed Type Stochastic Models for Population Growth", Theoretical Pop. Biol., 5, 358-365.

McNeil, D.R. (1977), "Interactive Data Analysis", John Wiley, London.

McNeil, W.H. (1977), "Plagues and Peoples", Blackwell, Oxford.

McNicoll, C. (1984), "Consequences of Rapid Population Growth: Overview and Assessment", Pop. and Development Rev., 10, 177-240.

Marchetti, C. (1985), "Swings, Cycles and the Global Economy", New Scientist, 106, No. 1454, 12-15.

Mar Molinero, C. (1984), "Trend Curves: The Specification Problem", presented at 4th International Symposium on Forecasting, London, 8-11 July.

Maston, R.G. (1981), "Comment on Hayden's 1981 Paper", Current Anthropology, 22, 537.

May, R.M. (1981), "Theoretical Ecology", 2nd ed., Blackwell Scientific, Oxford.

Meade, N. (1984), "The Use of Growth Curves in Forecasting Market Development - A Review and Appraisal", J. Forecasting, 3 429-451.

Meade, N. (1985), "Forecasting using Growth Curves - An Adaptive Approach", J. Operat. Res. Soc., 36, 1103-1116.

Metcalfe, J.S. (1981), "Impulse and Diffusion in the Study of Technical Change", Futures, 13, 347-359.

Minami, R. (1961), "An Analysis of Malthus' Population Theory", J. Econ. Behaviour, 1, 53-63.

Miner, J.R. (1932), "Birth Rate and Density in a Logistic Population", Human Biol. 4, 119-120.

Miner, J.R. (1933), "Pierre-Francois Verhulst - The Discoverer of the Logistic Curve", Human Biol., 5.

MINITAB 5.1 (1985), Minitab Inc., 3081 Enterprise Drive, State College, PA 16801, U.S.A.

Mirowski, P. (1976), "The Plague and the Penny-loaf: The Disease Dearth Nexus in Stuart and Hanoverian London", unpublished manuscript, Dept. of Economic, Univ. of Michigan.

Mitchell, B.R. (1980), "European Historical Statistics 1750-1975", Sec. Rev. Ed., Macmillan, London.

Montgomery, D.C. (1976), "Forecasting and Time Series Analysis", McGraw-Hill, New York.

Murphy, M.J. (1980), "Extrapolation of Current Trends for Forecasting Population", C.P.S. Working Paper 80-2, Centre of Population Studies, London School of Hygiene & Tropical Medicine, University of London.

Murphy, M.J. (1982), "Gompertz and Gompertz Relational Models for Forecasting Feritlity: An Empirical Exploration", C.P.S. Research Paper 82-2, Centre of Population Studies, London School of Hygiene & Tropical Medicine, University of London.

Murphy, M. (1984), "Forecasting Fertility in Developed Countries by the Use of ARIMA Models - An Exploratory Analysis", presented at the 4th International Symposium on Forecasting, 8-11 July, London.

Murphy, M.J. (1984), "Population Forecasting: Alternative Approaches", O.P.C.S. Occasional Paper.

Murphy, M.J. (1985), "Demographic and Socio-economic Influences on Recent British Marital Breakdown Patterns", Population Studies, 39, 441-460.

NAg 10 (1985), "The NAG FORTRAN Mark 10 Library", Numerical Algorithm Group Ltd., Oxford.

Nam, C.B. and Gustavus, S.O. (1976), "Population: The Dynamics of Demographic Change", Houghton Mifflin Company, Boston.

National Central Bureau of Statistics (1900-1981), "Statistisk Årsbok - Official Statistics of Sweden", The National Central Bureau of Statistics.

Nelder, J.A. (1961) "The Fitting of a Generalization of the Logistic Curve", Biometrics, 17, 89-110.

Norden, R.H. (1980), "An Analysis of a Stochastic Process of the Logistic Form", Bath Ph.D.

Nordic Statistics (1970-1978), "Yearbook of Nordic Statistics", Vols. 9-19, The Nordic Council and the Nordic Statistical Secretariat.

Notestein, F. (1983), "Population Growth and Economic Development", Pop. and Development Rev., 9, 345-360.

O.P.C.S. (Office of Population Censuses and Surveys) (1972), "Population Projections 1971-2011", No. 2, H.M.S.O., London.

O.P.C.S. (Office of Population Censuses and Surveys) (1984), "Population Projections 1981-2021", Ser. PP2, No. 12, H.M.S.O., London.

Ogden, P. (1975), "Empty France", Geographical Magazine, 94-97.

Oliver, F.R. (1964), "Methods of Estimating the Logistic Growth Function", Appl. Statist. 13, 57-66.

Oliver, F.R. (1966), "Aspects of Maximum Likelihood Estimation of the Logistic Growth Function", J. Amer. Statist. Assoc., 61, 697-705.

Oliver, F.R. (1969), "Another Generalisation of the Logistic Growth Function", Econometrica, 37, 144-147.

Oliver, F.R. (1981), "Tractors in Spain: A Further Logistic Analysis", J. Opl. Res. Soc. 32, 499-502.

Oliver, F.R. (1982), "Notes on the Logistic Curve for Human Populations", J.R. Statist. Soc., Ser. A, 145, 359-363.

Olney, M.L. (1983), "Fertility and the Standard of Living in Early Modern England: In Consideration of Wrigley and Schofield", J.E.C. Hist. 43, 71-77.

Oshima, H.T. (1983), "The Industrial and Demographic Transitions in East Asia", Pop. and Development Rev., 9, 583-607.

Parsons, J. (1977), "Population Fallacies", Elek/Pemberton, London.

Peacock, A.T. (1952), "Theory of Population and Modern Economic Analysis", Population Studies, 6, 114-122.

Pearl, R. (1925), "Studies in Human Biology", Williams and Wilkins Co., Baltimore.

Pearl, R. (1930), "The Biology of Population Growth", A.A. Knopf, New York.

Pearl, R. and Reed, L.J. (1920), "On the Rate of Growth of the Population of the United States since 1790 and its Mathematical Representation", from Proc. of the National Academy of Sciences, 6.

Pearl, R. and Reed, L.J. (1923), "On the Mathematical Theory of Population Growth", Merton, 3, 6-19.

Piatier, A. (1981), "Innovation Information and Long-term Growth", Futures, 13, 371-382.

Pielou, E.C. (1977), "Mathematical Ecology", Wiley, New York.

Pittenger, D.B. (1978), "The Role of Judgment, Assumptions, Techniques and Confidence Limits in Forecasting Population", Socio-Economic Planning Science, 12, 271-276.

Pittioni, R. (1981), "Comment on Hayden's 1981 Paper", Current Anthropology, 22, 538.

Plog, F. (1975), "Demographic Studies in South-western Prehistory", in Gumerman, 94-103.

Polgar, S. (1972), "Population History and Population Policies from an Anthropological Perspective", Current Anthropology, 13, 203-211.

Poston, T. and Stewart, I. (1978), "Catastrophe Theory and its Applications", Pitman, London.

Pritchett, A.S. (1891), "A Formula for Predicting the Population of the United States", Quart. Publ. Amer. Statist. Assoc., 2, 278-286.

Quetelet, A. (1835), "Sur l'Homme et le Developpement de ses Facultes ou Essai de Physique Sociale", Paris.

Raeside, R. (1984), "Documentation and Listing of PTRENDS program", Dept. of Mathematics, Napier College, Edinburgh.

Raeside, R. (1985), "Documentation and Listing of BIRTH program", Dept. of Mathematics, Napier College, Edinburgh.

Raeside, R. (1985), "Documentation and Listing of PKAL program", Dept. of Mathematics, Napier College, Edinburgh.

Ratkowsky, D.A. (1983), "Nonlinear Regression Modelling", Marcel Dekker Inc., New York.

Reed, H.S. and Holland, R.H. (1919), "The Growth Rate of an Annual Plant Helianthus", Proc. Nat. Acad. Sci. USA, 5, 35-144.

Reed, L.J. and Pearl, R. (1927), "On the Summation of Logistic Curves", J. Royal Statist. Soc., 90, 729-746.

Renfrew, C. (1978), "Trajectory Discontinuity and Morphogenesis: The Implications of Catastrophe Theory for Archaeology", Amer. Antiquity, 43, 203-222.

Renfrew, C. and Cooke, K.L. (eds.) (1979), "Transformations: Mathematical Approaches to Culture Change", Academic Press, London.

Renfrew, C.A. and Poston, T. (1979), "Discontinuities in the Endogenous Change of Settlement Pattern", in "Transformations: Mathematical Approaches to Culture Change", eds. K.L.Cooke and C.A. Renfrew, Academic Press, London, 437-461.

Repetto, R.and Holmes, T. (1983), "The Role of Population in Resource Depletion in Developing Countries", Pop. and

Development Rev., 9, 609-632.

Richards, F.J. (1959), "A Flexible Growth Function for Empirical Use", J. Experimental Botany, 10, 290-300.

Ripley, S. (1980), "Infanticide in Langurs and Man Adaptive Advantage or Social Pathology", in "Biosocial Mechanisms of Population Regulation", eds. Cohen et al., 349-390.

Robinson, P.M. (1985), ""Comment on Silverman's Paper", J.R. Statist. Soc., Ser. A, 47, 42.

Robinson, W.C. (1986), "High Fertility as Risk-Insurance", Population Studies, 40 289-298.

Ross, G.J.S. (1970), "The Efficient Use of Function Minimisation in Non-Linear Maximum-Likelihood Estimation", Appl. Statist., 19, 205-221.

Rostow, W.W. (1960), "The Process of Economic Growth", Clarendon Press, Oxford.

Rostow, W.W. (1980), "Why the Poor Get Richer and the Rich Slow Down", Macmillan, London.

Ryder, N. (1969), "The Emergence of a Modern Fertility Pattern; United States 1917-1966", in "Fertility and Family Planning: A World View", eds. S.J. Behrman et al., Univ. of Michigan Press, Ann. Arbor., 99-117.

Ryder, N.B. (1974), "The Family in Developed Countries", Scientif. Amer., 231, 123-132.

Saboia, J.L.M. (1974), "Modelling and Forecasting Populations by Time Series: The Swedish Case", Demography, 11, 483-492.

Saboia, J.L.M. (1977), "Autoregressive Integrated Moving Average (ARIMA) Models for Birth Forecasting", J. Amer. Statist. Assoc., 72, 264-270.

Samuelson, P.A. (1966), "Foundations of Economic Analysis", Harvard University Press, Cambridge, Massachusetts.

Sandland, R.L. and McGilchrist (1979), "Stochastic Growth Curve Analysis", Biometrics, 35, 255-271.

Schofield, R. (1983), "The Impact of Scarcity and Plenty on Population Change in England 1541-1871", J. Interdisciplinary History, 14, 265-291.

Schtickzelle, M. (1981), "Pierre-Francois Verhulst (1804-1849) La Premiere Decouverte de la Fonction Logistique", Population, 36, 541-556.

Schumacher, E.F. (1973), "Small is Beautiful", Blond and Briggs, London.

Schumpter, J.A. (1939), "Business Cycles", 2 vols. McGraw-Hill, London.

Schumpter, J. (1964), "Business Cycles", abridged version, McGraw-Hill, New York.

Schwefel, H.P. (1981), "Numerical Optimisation of Computer Models", John Wiley & Sons, Chichester.

Shryock, H.S. and Siegel, J.S. (1976), "The Method and Materials of Demography", Academic Press, London.

Silverman, B.W. (1985), "Some Aspects of the Spline Smoothing Approach to Non-Parametric Regression Curve Fitting", J.R. Statist. Soc., B, 47, 1-52.

Silverman, B.W. and Watters, G.W. (1984), "BATHSPLINE: An Interactive Spline Smoothing Package", University of Bath.

Simon, J.L. (1981), "The Ultimate Resource", Martin Robertson & Co., Oxford.

Simon, J.L. (1983), "The Present Value of Population Growth in the Western World", Population Studies, 37, 5-21.

Simons, J. (1977), "Developments in the Interpretation of Recent

Fertility Trends in England and Wales" in "Regional Demographic Development", eds. J. Hobcraft and P. Rees, Groom Helm, London.

Simons, J. (1986), "Culture Economy and Reproduction in Contemporary Society", in "The State of Population Theory", eds. D. Coleman and R. Schofield.

Smith D.P. (1982), "A Reconsideration of Easterlin Cycles", Population Studies, 35, 247-264.

Smith, P. (1972), "Changes in Population Pressure in Archaeological Explanation", World Archaeology, 4, 5-18.

Smith, R.M. (1981), "Fertility, Economy and Household Formation in England over Three Centuries", Pop. and Development Rev., 7, 595-622.

Smith, R.M. (1986), "Transfer Incomes, Risk and Security: The Roles of the Family and the Collectivity in Recent Theories of Fertility Change", in "The State of Population Theory", eds. D. Coleman and R. Schofield.

Spencer, P. (1973), "Nomads in Alliance: Symbiosis and Growth among the Rendille and Samburu of Kenya", Oxford University Press.

Spengler, J.J. (1972), "Population Economics", Duke Univ. Press, Durham, North Carolina.

Stone, R. (1979), "Sigmoids", Roy. Statist. Soc. Conference, Oxford, 2-6 April.

Stuart, A. (1986), "Who or What Killed the Giant Armadillo?", New Scientist, 111, 29-32.

Swedish Yearbook (1967-1976), "Statistisk Årsbok", Vols. 54-63, National Central Bureau of Statistics, Stockholm.

Tashchian, R.O. and Tashchian, A. (1984), "Logistic Curve

Parameter Estimation: Its Appalicatoin in Population Forecasting", 4th Int. Symposium on Forecasting, 8-11 July, London.

Thom, R. (1983), "Mathematical Models of Morphogenesis", Ellis Horwood Ltd., Chichester.

Thompson, D.A.W. (1942), "On Growth and Form", 2nd ed., Vol. 1, Cambridge University Press.

Thornton, H.G. (1922), "On the Development of a Standardised Agar Medium for Counting Soil Bacteria", Ann. Appl. Biol., 9, 241-274.

U.N. Dept. of International Economic and Social Affairs (1977), "Prospects of Population: Methodology and Assumptions", by Ad Hoc Group of Experts on Demographic Projections, ST/ESA/SER.A/67.

van Wyk, R.J. (1985), "The Notion of Technological Limits", Futures, 17, 215-223.

Verhulst, P.F. (1838), "Notice sur la Loi que la Population suti dans son Accroissement", Correspondence Mathematique et Physique, 10, 113-121.

Verhulst, P.F. (1847), "Deuxieme memoire sur l'accroissement de la population", Nouveaux Memoires de l'Academie Royale des Sciences et Belles Lettres de Bruxelles, 20, 1-32.

Vieira, S. and Hoffmann R. (1977), "Comparison of the Logistic and the Gompertz Growth Functoins considering Additive and Multiplicative Error Terms", Appl. Statist., 26, 143-148.

von Bertalanffy, L. (1938), "A Quantitative Theory of Organic Growth", Human Biol., 10, 181-213.

von Bertalanffy, L. (1949), "Problems of Organic Growth", Nature, 163, 156-158.

von Bertalanffy, L. (1957), "Quantitative Laws in Metabolism", Quart, Rev. Biol., 32, 217-231.

von Szeliski, V.S. (1936), "Population Growth due to Immigration and Natural Increase", Human Biol. 8, 25-37.

von Tunzelmann, G.N. (1979), "Trends in Real Wages, 1750-1850, Revisited", Ec. Hist. Rev., 23, 33-49.

von Tunzelmann, G.N. (1986), "Malthus and the Standard of Living Debate", in "The State of Population Theory", eds. D. Coleman and R. Schofield.

Wall, R. (1979), "Regional and Temporal Variations in English Household Structure from 1650", in "Regional Demographic Development", eds. J. Habcraft and P. Rees, Groom Helm, London.

Ward, M.P. and Butz, W.P. (1980) "Completed Fertility and its Timing", J. Political Economy, 88, 917-940.

Westoff, C.F. (1974) "The Populations of the Developed Countries", Scientific Amer., 231, 109-120.

Westoff, C.F. (1983), "Fertility Decline in the West: Courses and Prospects", Pop. and Development Rev., 9, 99-104.

Wichmann, B.A. and Hill I.D. (1985), "Algorithm AS183 - An Efficient and Portable Pseudo-Random Number Generator", in "Applied Statistics Algorithms", Ellis Horwood Ltd., Chichester.

Wilkinson, M. (1967), "Evidence of Long Swings in the Growth of Swedish Population and Related Economic Variables", J. Econ. History, 27, 17-38.

Willekens, F. (1984), "Population Forecasting", 4th Int. Symposium on Forecasting, London, 8-11 July.

Williams, J.E. (1966), "The British Standard of Living 1750-1850", Ec. Hist. Rev., 19, 581-589.

Wilmsen, E.N. (1975), "Population Pressure as a Non-Explanation", in Gumerman, 127-131.

Wilson, E.B. and Puffer, R.R. (1933), "Least Squares and Laws of Population Growth", Proc. Amer. Acad. Arts and Sciences, 68, 285-382.

Winsor, C.P. (1932), "The Gompertz Curve as a Growth Curve", Proc. National Academy of Sciences, 18, 1-8.

Winter, J.M. (1980), "The Fear of Population Decline in Western Europe 1870-1970", in "Demographic Patterns in Developed Societies", ed. R.W. Hiorns, Taylor and Francis Ltd., London, 173-197.

Wolfe, A.B. (1927), "Is there a Biological Law of Human Population Growth?", Quarterly J. Econ., 42, 557-594.

Woods, R.I. (1983), "On the Long-term Relationship between Fertility and the Standard of Living", Genus, 39, 21-35.

Wright, D.J. (1983), "Catastrophe Theory in Management Forecasting and Decision Making", Operat. Res., 34, 935-942.

Wright, F.J. (1982), "The Road-Crossing Catastrophe", The Mathematical Intelligencer, 4, 103-104.

Wrigley, E.A. (1969), "Population and History", McGraw-Hill, New York.

Wrigley, E.A. (1985), "The Fall of Marital Fertility in Nineteenth Century France", European J. of Pop., 1, 31-60.

Wrigley, E.A. and Scofield, R. (1981), "The Population History of England 1541-1871", Edward Arnold, London.

Wunsch, G.J. (1984), "Theories, Models and Knowledge: The Logic of Demographic Discovery", Genus, 40, 1-18.

Wynne-Edwards, V.C. (1962), "Animal Dispersion in Relation to Social Behaviour", Oliver and Boyd, Edinburgh.

Wynne-Edwards, V.C. (1978), "Intrinsic Population Control: An Introduction in Population Control by Social Behaviour", eds. F.J. Ebling & D.M. Stoddart, Inst. of Biol., London.

Yesner, D.R. (1981), "Comment on Hayden's 1981 paper", Current Anthropolgy, 22, 540.

Yiu, Y.Z. (1980), "The Prospects of Chinese Population Growth from 1978 to 2028", Renkou Yanjiu (Pop. Res.), 3, 18-25.

Young, P. and Ord, J.K. (1984), "The Use of Discounted Least Squares in Technological Forecasting", Working Paper No. 84-3, presented at ISF84, London, 8-11 July.

Yule, G.U. (1925), "The Growth of Population and the Factors which Control it", J.R. Statist. Soc., A, 88, 1-58.

Zahler, R.S. and Sussmann, H.J. (1977), "Claims and Accomplishments of Applied Catastrophe Theory", Nature, 269, 759-763.

Zambardino, R.A. (1980), "Mexico's Population in the Sixteenth Century: Demographic Anomaly or Mathematical Illusion?", J. Interdisplinary History, 11, 1-27.

Zeeman, E.C. (1976), "Catastrophe Theory", Scientif. Amer., 234, 65-83.

Zeeman, E.C. (1977), "Catastrophe Theory. Selected papers (1972-1977)", Addison-Wesley, Reading, Mass.

Zipf, G.K. (1949), "Human Behavior and the Principle of Least Effort", Addison-Wesley, Reading, Mass.