Reprinted from THEORETICAL POPULATION BIOLOGY
Vol. 20, No. 1, August 1981
Printed in Belgium

Dynamics of Single-Species Population Growth: Experimental and Statistical Analysis

LAURENCE D. MUELLER* AND FRANCISCO J. AYALA

Department of Genetics, University of California, Davis, California 95616

Received February 28, 1980

The logistic model, widely used for describing population growth, assumes that the per-capita rate of growth linearly decreases as the population size increases. Experimental data, however, suggest that often the per-capita rate of growth is not linearly related to population density. The theta model removes such linearity assumption by means of an additional parameter, θ; when $\theta = 1$, the theta model reduces to the logistic model. We advance a method, the "jackknife" statistic, for estimating the rate of population growth (the largest eigenvalue and its variance) in the serial transfer system. Also, we propose a statistical method, PRESS, for quantifying the success of a given model in fitting experimental data. The criterion of success is the ability of a model to predict accurately new observations. One advantage of PRESS is that, contrary to what happens with other statistics such as R^2, it tends to make a model less successful as the number of parameters increases (unless there is a disproportionate decrease in the bias of the new model). We have studied the rate of population growth in 25 genetically different populations of *Drosophila melanogaster*. The theta model provides a consistently better description of population growth in these populations than the logistic model. Moreover, the results indicate that the rate of growth is affected by the genetic constitution of a population.

1. INTRODUCTION

A widely used model describing population growth in a single species is the logistic model, proposed by Lotka (1924) and Volterra (1931). It predicts the rate of population growth as

$$dN/dt = rN(1 - N/K), \tag{1}$$

where N is the population size, r is the approximate per-capita rate of increase achieved at low densities, and K is the carrying capacity or equilibrium population size. The model predicts that the per-capita rate of increase ($N^{-1}dN/dt$) decreases from near r to 0 in a linear fashion as the population size increases; that is, the increase in intraspecific competition

* Present address: Department of Biological Sciences, Stanford University, Stanford, Calif. 94305.

0040-5809/81/040101-17$02.00/0

due to the addition of a new member is the same whether the total population size is small or large.

This linearity assumption can be removed in a variety of ways. For example, Schoener (1978) has produced some simple models of single-species population growth which include exploitative competition and interference competition. These models predict a faster-than-linear decline in per-capita rate of increase. Moreover, there is empirical evidence suggesting that the linearity assumption of the logistic model may often be violated (Smith, 1963; Ayala *et al.*, 1973).

More complicated versions of the logistic model are of course, possible, obtained by incorporating additional parameters into the logistic model (1). However, biologists are interested in simple models that capture the essential features of biological phenomena. Consequently, before the logistic model is abandoned, the benefits of more complicated models must be quantified in some meaningful way. Previous efforts have not always done this.

In this paper we ask whether relaxing the linearity assumption of the logistic model can yield a new model which can more accurately predict the dynamics of laboratory populations of *Drosophila melanogaster*. The robustness of the results is examined by appropriate statistical tests in a large number of independent populations. The experimental populations differ in their genetic constitution; hence, we also explore the possible contribution of genetic differences to population dynamics. A statistical test is described that determines whether the parameters estimated from the population growth equations are sensitive to the genetic constitution of populations. This is important because most theories of density-dependent selection assume that there is genetic variation in natural population for population parameters such as r and K of the logistic model (1).

2. Population Dynamics in a Serial Transfer System

2.1. *A General Model of the Serial Transfer System*

The Serial Transfer System (STS) of population growth is outlined in Fig. 1 (top). It is a discrete system of growth that allows for overlapping generations in the adult population. Two versions of the STS are possible, known as Type-1 and Type-2 experiments. In Type-1 experiments, the population grows until it reaches its carrying capacity, which is thereafter approximately maintained for the duration of the experiment. Type-2 experiments are used to determine the rate of population growth at a given density (Ayala *et al.*, 1973).

The procedure used in Type-1 experiments is described in Ayala (1965). Here, we shall describe in brief outline how the procedure is used with one-

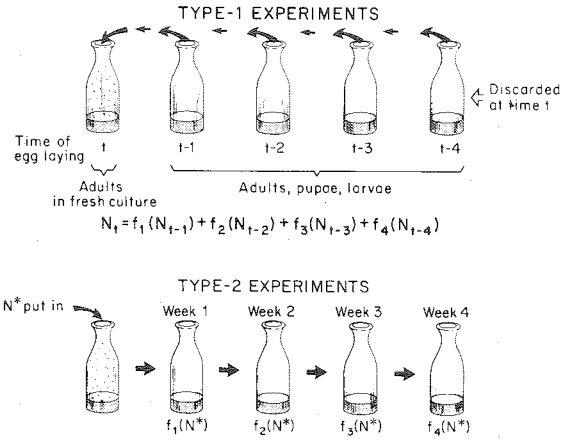

FIG. 1. The Serial Transfer System of population growth. Type-1 experiments are long-term populations. Type-2 experiments are used for estimating the rate of a population growth at a specified density, N^*.

week intervals between transfers. The population consists at any one time of four cultures of various ages. At the time of census, the youngest culture is one week old and contains egg-laying adults that have survived for one week since they were introduced in the culture. The remaining three cultures are 2, 3, and 4 weeks old and contain larvae, pupae, and newly emerged adults. At the time of census, the surviving adults in the youngest culture are counted and transferred to a fresh culture. At the same time, the newly emerged adults in the other three cultures are counted and are also transferred to the same fresh culture to which the adults were transferred; the oldest culture is, then, discarded. This process is summarized in Fig. 1 by arrows indicating transfer of adults from the four old cultures to the one fresh culture. The adults introduced in the fresh culture are allowed to lay eggs for one week, when a new census is taken. The procedure is repeated at one-week intervals.

At the time of census, the total number of adults, N_t, is a function of the number of adults present in the leading culture 1, 2, 3, and 4 weeks ago. This relationship can be summarized as

$$N_t = f_1(N_{t-1}) + f_2(N_{t-2}) + f_3(N_{t-3}) + f_4(N_{t-4}), \tag{2}$$

where $f_i(N_{t-i})$ is some unknown function that relates the number of adults that will emerge from an i-weeks-old culture with the number of adults that initially laid eggs in that culture. The exact form of these functions is of little general interest; as we shall show below, the analysis of population growth rates can proceed by an appropriate linearization of this model.

2.2. *Experimental Determination of Rates of Population Growth at a Single Density*

In order to investigate the usefulness of models such as (1) we need to obtain estimates of population growth rates at a variety of densities. This sort of information can be obtained by STS Type-2 experiments as shown in Fig. 1 (bottom) (Ayala *et al.*, 1973). A specified number of adults, N^*, are initially placed in a fresh culture. The suvivors are counted one week later and the emerging adults from this same culture over the following 3 weeks are recorded at one-week intervals. Each experiment of this kind yields the vector $(Y_1(N^*), Y_2(N^*), Y_3(N^*), Y_4(N^*))$, where $Y_i(N^*) = f_i(N^*) + \varepsilon_i$, and ε_i is a random variable reflecting experimental error. Thus $Y_i(N^*)$ is the observed number of adults emerging (or surviving in the case of $Y_1(N^*)$) from an i-weeks old culture of egg laying by N adults during one week. Type-2 experiments can be repeated for multiple independent estimates of a given $f_i(N^*)$ and can be carried out at various densities.

In order to estimate the rate of population growth in the vicinity of N^*, we look at a linear version of (2),

$$N_t = a_1 N_{t-1} + a_2 N_{t-2} + a_3 N_{t-3} + a_4 N_{t-4}, \tag{3}$$

where each a_i is a constant per-capita output of an i-weeks-old culture. Equation (3) is a fourth order homogeneous difference equation. In the case when all four eigenvalues of this equation are real and distinct, the general solution to (3) may be written as

$$N_t = c_1 \lambda_1^t + c_2 \lambda_2^t + c_3 \lambda_3^t + c_4 \lambda_4^t, \tag{4}$$

where the c's are constants which can only be determined if the four initial conditions are specified. Unfortunately there are no obvious initial conditions that can be used to obtain an explicit solution to (3). Moreover, these experiments are done at a variety of densities and it does not seem reasonable that the initial conditions used at one density will be compatible with the initial conditions at any other density. However, as t gets large in (4) a per-capita rate of population growth is obtained that is independent of the initial conditions. In such case the approximate per-capita rate of increase will be given by the one positive eigenvalue of (3), $N_t/N_{t-1} \simeq \lambda$. It is then possible to determine the weekly change in population size as a function of N^*, as $\Delta N/\Delta t(N^*) = (\lambda N^*) - N^* = N^*(\lambda - 1)$.

Smith (1963) encountered a similar situation; i.e., where the initial conditions determine the outcome. In his experiments growing populations of *Daphnia* were maintained until a stable age distribution was attained. The relationship between density and rate of population growth was sensitive to the initial condition used to start the experiment. Repeatable results were obtained only when populations had attained a stable age distribution.

2.3. *Statistical Estimation of Rate of Population Growth*

If the Type-2 experiments (Section 2.2) are carried out m times at a density N^*, then these m observations, $[Y_1^{(1)}(N^*), Y_2^{(1)}(N^*), Y_3^{(1)}(N^*), Y_4^{(1)}(N^*)],\ldots, [Y_1^{(m)}(N^*), Y_2^{(m)}(N^*), Y_3^{(m)}(N^*), Y_4^{(m)}(N^*)]$, can be used to estimate the a's in Eq. (3). Our procedure will be first to use all the observations to estimate the a's in (3) and then determine the largest eigenvalue of the resulting equation. It is also possible to estimate the largest eigenvalue for each observation and take the average of these m values. The results obtained by these two methods need not be the same. Given that the eigenvalues of (3) are functions of the population quantities a_i, it seems more appropriate to estimate first the a_i's and then use these for determining the eigenvalues.

Therefore, we will estimate each a_i by $\hat{a}_i = 1/m \sum_{j=1}^{m} Y_i^{(j)}/N^*$, $i = 1, 2, 3, 4$, where the argument N^* has been deleted for simplicity. This yields one difference equation, $N_t = \hat{a}_1 N_{t-1} + \hat{a}_2 N_{t-2} + \hat{a}_3 N_{t-3} + \hat{a}_4 N_{t-4}$, from which an estimate of the largest eigenvalue $\hat{\lambda}$ is obtained. There is, unfortunately, no simple way to estimate the variance of $\hat{\lambda}$. One approximate solution is the jackknife statistic (for a review see Miller, 1974; for applications to population genetics see Mueller, 1979).

In order to calculate the jackknife statistic, the jth set of observations is deleted and the largest eigenvalue (as described in the previous paragraph) is calculated using the remaining data. This yields a new value $\hat{\lambda}_{-j}$. One can then calculate m pseudovalues as

$$s_j = m\hat{\lambda} - (m-1)\hat{\lambda}_{-j}, \qquad j = 1, 2,\ldots, m.$$

The jackknifed estimate of the largest eigenvalue is simply the mean of these pseudovalues; $\tilde{\lambda} = (1/m) \sum_j s_j$. The variance of this eigenvalue is estimated by $\hat{\text{Var}}(\tilde{\lambda}) = (1/m(m-1)) \sum_j (s_j - \tilde{\lambda})^2$. The pseudovalues may also be used to estimate m values of $\Delta N/\Delta t$ as $N^*(s_j - 1)$, $j = 1, 2,\ldots, m$. These values of $\Delta N/\Delta t$ are necessary for the regression analysis described in the next section.

3. Statistical Methods

3.1. *PRESS*

One of our aims is to compare different models that predict the change in population size as a function of density. As mentioned in the Introduction, it is important to quantify the success of each model in some meaningful way. Two widely used statistics that measure how well a regression function fits a set of data are the proportion of variance explained by the model (R^2) and the mean sum of squares ($1/n$ RSS). R^2 will increase and ($1/n$) RSS will

decrease as more complex elaborations of some basic model are examined. In the limit, if there are $n+1$ observations it is always possible to obtain an nth order polynomial which passes exactly through all $n+1$ points. This polynomial yields $R^2 = 1$ and $(1/n)\text{RSS} = 0$. Given such direct correlation between the number of parameters and goodness of fit, it is not clear what the best model might be. It is obviously necessary to establish criteria for deciding which model is best.

For the present purposes, we consider the best model that model that can predict new observations most accurately. The statistic PRESS (Prediction Sum of Squares; Allen, 1971a) provides a means for quantifying this property. We will describe this more formally. Consider linear models where $\boldsymbol{\beta}$ is the vector of parameters of the regression function (e.g., r and K of (1)). Similar arguments can be made for nonlinear models also. Suppose we employ the following partitioning of $\boldsymbol{\beta} = (\beta_1, \beta_2)^T$. Now let $\hat{\boldsymbol{\beta}}$ be the least-squares estimate of $\boldsymbol{\beta}$ and $\tilde{\boldsymbol{\beta}}$ be the least-squares under the assumption $\beta_2 = 0$. If $\boldsymbol{\lambda}$ is a vector of independent variables then we have two estimators of $\boldsymbol{\lambda}^T\boldsymbol{\beta}$ available: $\boldsymbol{\lambda}^T\hat{\boldsymbol{\beta}}$ and $\boldsymbol{\lambda}^T\tilde{\boldsymbol{\beta}}$. It can be shown (Walls and Weeks, 1969) that $\text{Var}(\boldsymbol{\lambda}^T\hat{\boldsymbol{\beta}}) \geqslant \text{Var}(\boldsymbol{\lambda}^T\tilde{\boldsymbol{\beta}})$ whether or not $\beta_2 = 0$. Obviously if $\beta_2 \neq 0$ then $\boldsymbol{\lambda}^T\tilde{\boldsymbol{\beta}}$ will be biased. Thus the following dilemma exists: the use of $\boldsymbol{\lambda}^T\tilde{\boldsymbol{\beta}}$ as an estimate of $\boldsymbol{\lambda}^T\boldsymbol{\beta}$ will have a small variance but may be biased, while use of $\boldsymbol{\lambda}^T\hat{\boldsymbol{\beta}}$ will reduce this bias at a cost of increased variance. The problem is to strike a balance between variance and bias. In such a circumstance it may be best to use the estimator with the smallest mean-squarred error since this equals the variance plus the bias squared. PRESS provides an estimate of the mean-squared error. Let $g(N, \hat{\boldsymbol{\beta}})$ represent the regression function, where N is the observed density and $\hat{\boldsymbol{\beta}}$ is the vector of estimated parameters; then, PRESS is defined for a sample of n observations as

$$\text{PRESS} = \frac{1}{n}\sum_{i=1}^{n}[U^{(i)} - g(N^{(i)}, \hat{\boldsymbol{\beta}}^{(-i)})]^2,$$

where $U^{(i)}$ is the observed value of $\Delta N/\Delta t(N^{(i)})$ and $\hat{\boldsymbol{\beta}}^{(-i)}$ has been estimated from the $n-1$ observations $(U^{(1)}, N^{(1)}),\ldots, (U^{(i-1)}, N^{(i-1)}), (U^{(i+1)}, N^{(i+1)}),\ldots, (U^{(n)}, N^{(n)})$. According to the present criterion, the best model is the one yielding the smallest value of PRESS.

3.2. *Variation for Regression Parameters*

Suppose a regression model has two parameters which must be estimated; that is, $\hat{\boldsymbol{\beta}} = (\hat{\beta}_1, \hat{\beta}_2)$. For model (1), β_1 corresponds to r and β_2 to K. In addition, there are l independent populations, for which $\hat{\boldsymbol{\beta}}$ (i.e., $\hat{\boldsymbol{\beta}}^{(1)},\ldots,\hat{\boldsymbol{\beta}}^{(l)}$) has been estimated. The question arises whether the l values of $\hat{\beta}_1$ and $\hat{\beta}_2$ show significant heterogeneity. In our study, the l populations represent

populations that are genetically different (each is homozygous for a different second chromosome). If we estimate r and K from Eq. (1) for each of these populations, the question is whether the genetic constitution of these populations has a significant effect on the estimated values of r and K. For linear models such as (1), an analysis of covariance can be used to answer this question. Analogous methods do not exist for nonlinear models. Since we shall be considering, in addition to the logistic model, a nonlinear model of population growth, we will adopt a method that is applicable to both linear and nonlinear models.

Suppose we examine the variation of parameter $\hat{\beta}_1$ over the l populations. The arguments that follow are unchanged when parameters other than $\hat{\beta}_1$ are considered; hence, we will drop the subscript 1 in the following discussion. Let the vector of these l parameters be $\hat{\boldsymbol{\beta}} = (\hat{\beta}^{(1)},\ldots,\hat{\beta}^{(l)})^T$. For each $\hat{\beta}^{(i)}$ we also have an estimated variance $\hat{\sigma}_i^2$. We assume that $\hat{\boldsymbol{\beta}}$ has an multivariate normal distribution with parameters $(\boldsymbol{\beta}, \Sigma)$, where $\Sigma = \text{diag}(\sigma_1^2,\ldots,\sigma_l^2)$. Let $\hat{\alpha}$ be some linear combination of the $\hat{\beta}$'s, $\hat{\alpha} = \sum_{i=1}^{l} h_i\hat{\beta}^{(i)} = \mathbf{h}^T\hat{\boldsymbol{\beta}}$. Using the methods of Gold (1963) and Goodman (1964), it can be shown that for all $\mathbf{h}^T$, $P[\hat{\alpha} - L \leqslant \alpha \leqslant \hat{\alpha} + L] \geqslant \gamma$, where $L = \sqrt{\text{Var}(\hat{\alpha})c}$ and $P(\chi_l^2 < c) = \gamma$. Thus this method generates $100\gamma\%$ simultaneous confidence intervals.

To determine whether there are significant differences between the $\hat{\beta}$ values, the order statistics of the $\hat{\beta}$, have been divided below into three groups. The choice of three is totally arbitrary. If a 95% simultaneous confidence interval on the difference between the means of any of these groups does not include 0, we conclude the $\hat{\beta}$'s are not homogeneous. This simultaneous inference scheme allows one to "hunt" for contrasts among the $\hat{\beta}$'s that will yield significant differences and still claim that this inference is being made at the $100\gamma\%$ confidence level.

4. The Models

In addition to the logistic model we will estimate the parameters and calculate PRESS for the theta model (Ayala *et al.*, 1973; Gilpin and Ayala, 1973):

$$dN/dt = rN[1 - (N/K)^{\theta}]. \tag{5}$$

There are several reasons for examining this model. Previous results (Ayala *et al.*, 1973) indicate that interspecific competition can be modeled best, when using the value of R^2 as a criterion, with analogs of (5). Moreover, the theta model includes the logistic model as a special case, i.e., when $\theta = 1$. Since our major interest is in relaxing the linearity assumption of (1), the theta model is a good option, because it allows for inflection points both less

than and greater than $K/2$. The recent models of Schoener (1978) have inflection points always less than $K/2$.

Only one linear regression need be performed in order to estimate PRESS (Allen, 1971b) for linear models such as (1). Nonlinear models, however, require repeated estimation of $\hat{\boldsymbol{\beta}}^{(-i)}$. If each population has n observations, then calculation of PRESS for one population requires n iterations of the nonlinear estimation procedure. This is prohibitively time consuming. The number of computations can be cut in half by deleting two observations, $(U^{(i)}, N^{(i)})$, $(U^{(j)}, N^{(j)})$, at a time; estimating the new vector of parameters $\hat{\boldsymbol{\beta}}^{(-ij)}$ on this set of $n-2$ observations; and then computing $[U^{(i)} - g(N^{(i)}, \hat{\boldsymbol{\beta}}^{(-ij)})]^2$, $[U^{(j)} - g(N^{(j)}, \hat{\boldsymbol{\beta}}^{(-ij)})]^2$. The bias, if any, introduced by this procedure would be to generate larger values of PRESS simply because less information is available for estimating $\hat{\boldsymbol{\beta}}^{(-ij)}$ than for estimating $\hat{\boldsymbol{\beta}}^{(-i)}$. This bias will only be present in estimates of PRESS for the theta model. Despite this bias, the theta model has a smaller value of PRESS most often and, hence, our conclusions will be unaffected by this bias.

In this study, the parameters of (5) are estimated using the algorithm of Marquardt (1963). Marquardt's algorithm uses a ridge-regression improvement at each iteration of the algorithm. This is particularly useful when there is a high degree of correlation between the parameters. Problems that occur when parameters are highly correlated are (a) round-off errors in the numerical procedure used may lead to inaccurate results; and (b) the estimated parameters have a very large variance. Ridge estimators are useful in combating these problems (Marquardt and Snee, 1975).

The results presented in Section 6 show that r and θ are negatively correlated. This correlation is a consequence of the following. At each step of the nonlinear routine, the following linear regression problem is solved (Gallant, 1975),

$$\mathbf{z} = \mathbf{x}\boldsymbol{\beta},$$

where

$$\mathbf{z} = \mathbf{U} - \mathbf{g}(\mathbf{N}, \hat{\boldsymbol{\beta}}_t) + \mathbf{x}\hat{\boldsymbol{\beta}}_t,$$
$$\mathbf{U} = (U^{(1)}, U^{(2)}, \ldots, U^{(n)})^T,$$
$$\mathbf{g}(\mathbf{N}, \hat{\boldsymbol{\beta}}_t) = (g(N^{(1)}, \hat{\boldsymbol{\beta}}_t), \ldots, g(N^{(n)}, \hat{\boldsymbol{\beta}}_t))^T,$$

At the tth iteration,

$$\hat{\boldsymbol{\beta}}_t = (\hat{\beta}_1, \ldots, \hat{\beta}_P),$$
$$x = \{x_{ij}\} = \left\{ \frac{\partial}{\partial \beta^{(j)}} g(N, \boldsymbol{\beta}) \Big|_{\boldsymbol{\beta} = \hat{\boldsymbol{\beta}}_t}^{N = N^{(i)}} \right\}.$$

The next value of $\hat{\boldsymbol{\beta}}$ is given by the standard solution to the above problem, $\hat{\boldsymbol{\beta}}_{t+1} = (\mathbf{x}^T\mathbf{x})^{-1}\mathbf{x}^T\mathbf{z}$. For model (5), $\hat{\boldsymbol{\beta}} = (\hat{r}, \hat{K}, \hat{\theta})^T$; it is easy to show, in the above notation, that

$$x_{i3} = (x_{i1} - N_i)(r \ln N_i - r \ln K).$$

This close correspondence between entries in the first and third columns of $\mathbf{x}$ causes r and θ to be highly correlated.

5. Experimental Tests

Drosophila melanogaster flies were collected at Strawberry Canyon, Berkeley, California. Crosses of individual wild males with males with balancer stocks produced a number of lines, each homozygous for the complete second chromosome (see Tracey and Ayala, 1974). A total of 24 nonlethal and nonsterile lines were selected for Type-2 experiments in order to measure the rate of population growth in each line. Type-2 experiments were also done with a random-heterozygous line (H) used as a standard or reference. The H line was produced by placing five virgin females and five males from each homozygous line in each of 10 cultures. These flies were allowed to mate at random and lay eggs for the next 5 days. F_1 progenies emerging from these cultures were placed in fresh cultures at the various densities (10, 50, 100, 250, 500, 750, and 1000) used in the experiments. The F_2 progenies produced in these cultures were used to start Type-2 experiments. This procedure was repeated three times; i.e., each time that a new set of replicates was started with the homozygous lines.

As described in Section 2.2, N^* adult flies were used to start each experiment. For each line, N^* took on values of 10, 20, 50, 100, 250, 500, 750, and 1000. Six replicates were made for each homozygous line at each density, except at 1000 for which only three replicates were done. Exceptions were line 45, which had only two replicates at density 1000; and line 36, which had no observations at 1000 and only three at 750. The H line was replicated 12 times at each density, except 1000 at which only three replicates were made. Two replicates of each homozygous line (and four of the H line) were started simultaneously, so that a total of three sets of replicates were started at different times. The variance between replicates started at different times was always much greater than the variance between replicates started at the same time. Most of this increased variance is probably due to differences between batches of the culture medium, although variations in incubator temperature and other environmental variables may have also contributed to this variance.

All experiments were performed in 237 cc (half-pint) cultures with 40 cc of

a standard cornmeal–molasses–agar medium. The cultures were kept at 23°C and ca. 70% relative humidity. All adults were between 7 and 14 days old at the start of each experiment, and had been raised at the same density condition as the density used in the experiment. An equal number of males and females were placed in each culture, in order to standardize the experimental conditions over all lines.

6. Results

Tables I and II give the estimated values of the parameters for the logistic (1) and the theta (5) model for each homozygous line as well as for the random heterozygous (H) line. Standard errors and correlation coefficients

TABLE I

Values of the Parameters (with Their Standard Errors) of the Logistic Model, Estimated for Each of 25 Genetically Different Lines

Line	r	K	Correlation coefficient between r and K
H	1.01 ± 0.08	720 ± 16	0.26
1	0.99 ± 0.07	850 ± 17	−0.07
2	0.67 ± 0.07	850 ± 25	−0.09
3	0.54 ± 0.10	750 ± 37	−0.33
6	1.07 ± 0.10	1000 ± 40	−0.50
7	0.59 ± 0.06	670 ± 21	−0.64
8	0.96 ± 0.06	880 ± 18	−0.21
9	0.85 ± 0.08	900 ± 27	−0.25
13	0.76 ± 0.05	680 ± 14	0.60
14	0.97 ± 0.06	970 ± 24	−0.44
15	1.09 ± 0.07	770 ± 13	0.27
18	1.19 ± 0.06	1000 ± 21	−0.50
20	1.11 ± 0.08	900 ± 20	−0.26
23	0.81 ± 0.09	740 ± 24	0.38
25	1.11 ± 0.10	810 ± 19	0.09
30	0.94 ± 0.09	620 ± 19	0.75
33	0.85 ± 0.10	770 ± 26	0.24
36	0.72 ± 0.34	650 ± 10	−0.08
37	1.04 ± 0.09	710 ± 18	0.52
40	0.77 ± 0.08	760 ± 21	0.31
42	0.75 ± 0.08	940 ± 33	−0.36
43	0.64 ± 0.09	660 ± 28	0.66
45	0.79 ± 0.08	660 ± 20	0.57
50	1.01 ± 0.05	820 ± 12	0.03
52	0.48 ± 0.10	530 ± 45	0.89

TABLE II

Values of the Parameters (with Their Standard Errors) of the Theta Model, Estimated for Each of 25 Genetically Different Lines

Line	r	K	θ	Correlation coefficients $r\theta$	rK	$K\theta$
H	370 ± 5	680 ± 13	0.0022 ± 0.0001	−0.61	−0.13	0.06
1	140 ± 7	830 ± 20	0.0047 ± 0.0004	−0.61	0.07	−0.31
2	170 ± 7	830 ± 22	0.0029 ± 0.0002	−0.61	0.07	−0.31
3	80 ± 12	720 ± 30	0.0064 ± 0.0008	−0.60	−0.14	0.10
6	150 ± 9	1100 ± 53	0.0045 ± 0.0005	−0.63	−0.33	−0.71
7	44 ± 11	610 ± 24	0.0108 ± 0.0012	−0.58	−0.32	0.49
8	150 ± 6	880 ± 21	0.0044 ± 0.0003	−0.61	0.15	−0.44
9	250 ± 9	900 ± 34	0.0022 ± 0.0002	−0.62	0.18	−0.49
13	52 ± 8	610 ± 19	0.0113 ± 0.0009	−0.58	−0.33	−0.50
14	210 ± 4	1000 ± 20	0.0029 ± 0.0001	−0.62	0.29	−0.66
15	130 ± 6	720 ± 15	0.0062 ± 0.0004	−0.60	−0.13	0.08
18	140 ± 4	1100 ± 24	0.0050 ± 0.0002	−0.63	0.35	−0.74
20	99 ± 6	910 ± 24	0.0073 ± 0.0005	−0.62	0.19	−0.50
23	100 ± 12	690 ± 27	0.0063 ± 0.0008	−0.60	−0.20	0.21
25	190 ± 7	770 ± 20	0.0042 ± 0.0003	−0.61	−0.04	−0.12
30	57 ± 18	550 ± 41	0.0109 ± 0.0019	−0.57	−0.38	0.66
33	180 ± 7	720 ± 18	0.0040 ± 0.0003	−0.60	−0.13	0.07
36	65 ± 5	645 ± 15	0.0070 ± 0.0005	−0.66	0.09	−0.28
37	140 ± 7	652 ± 15	0.0063 ± 0.0004	−0.59	−0.27	0.35
40	99 ± 9	713 ± 21	0.0060 ± 0.0005	−0.60	−0.16	0.17
42	82 ± 10	955 ± 41	0.0062 ± 0.0007	−0.62	0.24	−0.59
43	460 ± 9	610 ± 20	0.0013 ± 0.0001	−0.59	−0.32	0.49
45	54 ± 9	604 ± 21	0.0117 ± 0.0011	−0.59	−0.30	0.43
50	130 ± 5	795 ± 13	0.0054 ± 0.0003	−0.61	0.005	−0.19
52	710 ± 11	502 ± 24	0.0009 ± 0.0008	−0.57	−0.41	0.74

are also included in these tables. As mentioned earlier, the strong negative correlation observed for the theta model between r and θ is expected in view of the design matrix used in the nonlinear analysis. It should also be noted that there is, in general, good agreement between the estimates of K from models (1) and (5), but that the values of r are drastically different.

Large differences between lines with respect to the parameter values are apparent in Tables I and II. Multiple contrasts between groups of homozygous lines obtained by the methods described in Section 3.2, are shown in Figs. 2, 3, and 4. Significant differences between homozygous lines exist for all parameters in both models. We believe that these differences are real phenomena, largely due to genetic differences among the populations.

Table III shows the values of PRESS for each line according to each

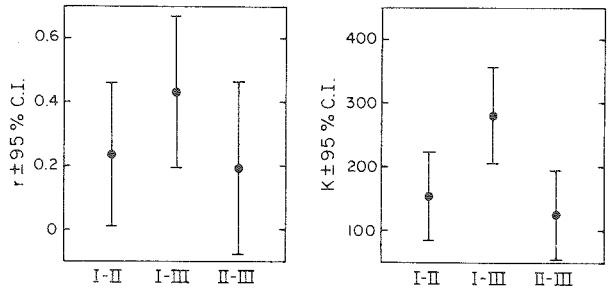

FIG. 2. Mean difference (with 95% confidence interval) between groups of populations for the parameters r (left) and K (right) of the logistic model. In the case of r, the populations involved in the multiple contrasts are I = 1, 6, 8, 14, 15, 18, 20, 25, 30, 37, 50; II = 9, 13, 23, 33, 40, 42, 45; III = 2, 3, 7, 36, 43, 52. In the case of K, the multiple contrasts are between I = 1, 2, 6, 8, 9, 14, 18, 20, 42; II = 3, 15, 23, 25, 33, 37, 40, 50; III = 7, 13, 30, 36, 43, 45, 52.

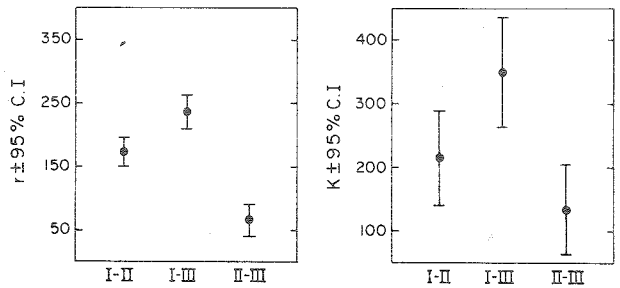

FIG. 3. Mean difference (with 95% confidence interval) between groups of populations for the parameters r (left) and K (right) of the theta model. In the case of r, the populations involved in the multiple contrasts are I = 2, 9, 14, 25, 33, 43, 52; II = 1, 6, 8, 15, 18, 37, 50; III = 3, 7, 13, 20, 23, 30, 36, 40, 42, 45. In the case of K, the populations compared are the same as for K in Fig. 2.

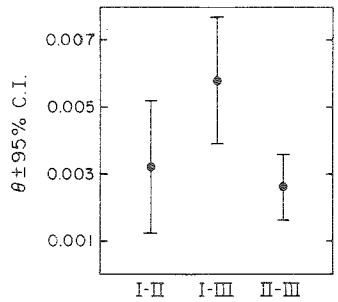

FIG. 4. Mean difference (with 95% confidence interval) between groups of populations for the parameter θ of the theta model. The populations in each group are the same as for r in Fig. 3.

TABLE III

PRESS Values (in Units of 10^3) for the Logistic and Theta Models in Each of 25 Genetically Different Lines

Line	Logistic	Theta
H	7.6	3.7
1	3.7	2.4
2	3.0	1.4
3	7.2	4.6
6	8.4	5.7
7	3.5	2.6
8	3.0	1.8
9	4.4	3.0
13	2.1	3.1
14	3.0	1.0
15	3.4	2.7
18	2.5	0.9
20	4.1	2.7
23	6.4	5.7
25	6.2	4.1
30	5.8	17
33	7.4	4.2
36	0.4	2.3
37	6.7	3.9
40	4.2	3.1
42	4.1	2.9
43	5.7	3.0
45	4.8	4.0
50	2.0	1.2
52	7.2	4.2

model. The results are quite clear. The theta model yields a lower value of PRESS than the logistic model for 22 out of the 25 populations. Although the difference in performance between the two models is unambiguous, the study of a single line (one of the three for which the logistic model gives a lower value of PRESS) might have led to erroneous conclusions. Figs. 5, 6 and 7 show the experimental data along with the predictions derived from the logistic (1) and the theta (5) model for three homozygous lines (18, 33, and 36); the theta model fits the observations conspicuously better.

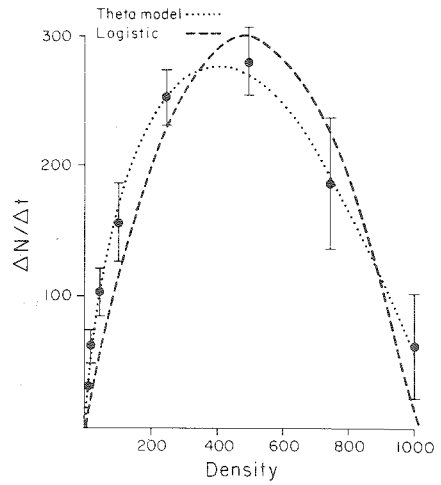

FIG. 5. $\Delta N/\Delta t$ (means with 95% confidence intervals) as a function of density in line 18. The theta model (dot line) fits the experimental data better than the logistic model (dash line).

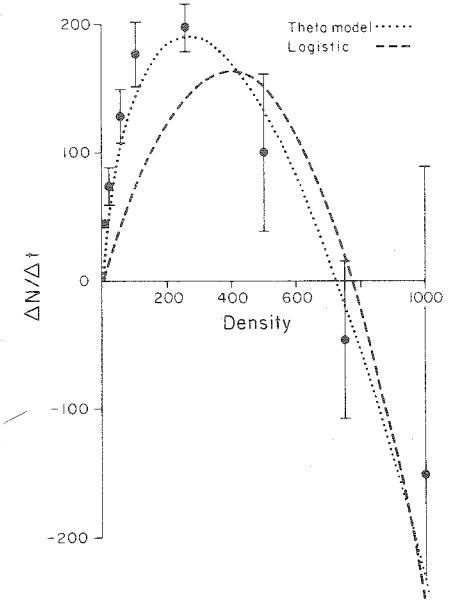

FIG. 6. $\Delta N/\Delta t$ (means with 95% confidence intervals) as a function of density in line 33. The theta model (dot line) fits the experimental data better than the logistic model (dash line).

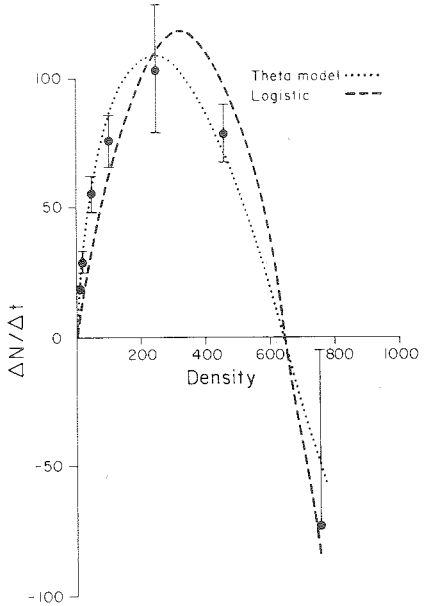

FIG. 7. $\Delta N/\Delta t$ (means with 95% confidence intervals) as a function of density in line 36. The theta model (dot line) fits the experimental data better than the logistic model (dash line) although PRESS is smaller for the logistic model.

7. DISCUSSION

The results presented in Section 6 show that relaxing the linearity assumption of the logistic model (1), as it is done by Eq. (5), yields a model that predicts new observations more accurately. The generality of this conclusion remains unknown, although nonlinear departures from the logistic model may be a common phenomenon (e.g., Smith, 1963; Ayala *et al.*, 1973; Schoener, 1973; Thomas *et al.*, 1980; Serradilla, 1979).

The inflection points dN/dt relative to N in all our experimental populations are at value of N smaller than $K/2$. Other detailed studies of population dynamics in *Daphnia* (Smith, 1963) and *Drosophila* (Thomas *et al.*, 1980; Seradilla, 1979) yield similar results. Schoener (1973) has made a crude estimate of the position of the inflection point for various species and has concluded that inflection points greater than $K/2$ are more common. His analysis includes two *Drosophila* populations with estimated inflection points greater than $K/2$. Given the high level of approximation used in Schoener's calculations, it would seem that a general conclusion cannot be reached until more detailed analyses, such as done in the present study, are available.

The desirable properties of mathematical models include generality and precision (Levins, 1966) as well as simplicity—i.e., having only the

minimum number of necessary parameters (Ayala *et al.*, 1973). Models should also have realism; that is, biological interpretation, in the case of population ecology. Hence, we are interested not only in a functional relationship, such as expressed in (5), that may describe "best" the nonlinearities observed in the experiments, but also in the biological phenomena responsible for the nonlinearities. Uncovering the underlying biological processes would make it possible to assess the applicability of models such as (5) to other organisms and to other populations. At present, only speculations are available. Schoener (1978) has shown that nonlinearity may simply result from the mechanism of feeding for a limited food supply. Gilpin *et al.* (1976) have suggested that nonlinearity may arise because quality resources are exhausted first in an environment with heterogeneous resources. A decision between these and other possible explanations must wait until experimental tests are performed addressed to ascertain the biological processes that account for nonlinearity in intraspecific competition.

There is one major reason why the values of r estimated from the theta and from the logistic model are quite different. In the logistic model, r determines the approximate rate of population growth at low densities. In the theta model, the rate of population growth is determined by the parameter $r\theta$. When θ is close to 0, the first two terms of the Taylor series expansion of (5) yield $\Delta N/\Delta t \simeq r\theta N \ln(K/N)$. Thus if one compares $r\theta$ values estimated from the theta model and r values from the logistic, one sees similar predictions of the growth rate of the populations at low densities.

This study has demonstrated that the growth dynamics of single-species populations are sensitive to the genetic constitution of populations. We plan to explore in future papers the relationships between genetic variation, Darwinian fitness, and density-dependent selection.

Acknowledgments

We would like to thank T. Bierbaum, M. Feldman, T. Prout and M. Turelli for valuable comments. This work was supported by Contract PA 200-14 Mod No. 4 with the U. S. Department of Energy and by NIH Grant P01 GM22221.

References

Allen, D. M. 1971a. Mean square error of prediction as a criterion for selecting variables, *Technometrics* **13**, 469–475.

Allen, D. M. 1971b. "The Prediction Sum of Squares as a Criterion for Selecting Predictor Variables," University of Kentucky Technical Report No. 23.

Ayala, F. J. 1965. Relative fitness of populations of Drosophila *serrata* and *Drosophila birchii*, *Genetics* **51**, 527–544.

Ayala, F. J., Gilpin, M. E., and Ehrenfield, J. G. 1973. Competition between species: Theoretical models and experimental tests, *Theoret. Population Biol.* **4**, 331–356.

Gallant, A. R. 1975. Nonlinear regression, *Amer. Statistician* **29**, 73–81.

Gilpin, M. E., and Ayala, F. J. 1973. Global models of growth and competition, *Proc. Nat. Acad. Sci.* **70**, 3590–3595.

Gilpin, M. E., Case, T. J., and Ayala, F. J. 1976. θ-selection, *Math. Biosci.* **32**, 131–139.

Gold, R. Z. 1963. Tests auxiliary to χ^2 tests in a Markov chain. *Ann. Math. Statist.* **34**, 56–74.

Goodman, L. A. 1964. Interactions in multidimensional contingency tables. *Ann. Math. Statist.* **35**, 632–646.

Lotka, A. J. 1924. "Elements of Mathematical Biology," reprinted 1956, Dover, New York.

Marquard, D. W. 1963. An algorithm for least-squares estimation of nonlinear parameters, *Soc. Indust, Appl. Math.* **11**, 431–441.

Marquard, D. W., and Snee, R. D. 1975. Ridge regression in practice, *Amer. Statistician* **29**, 3–20.

Miller, R. G. 1974. The jackknife—a review, *Biometrika* **61**, 1–15.

Mueller, L. D. 1979. A comparison of two methods for making statistical inferences on Nei's measure of genetic distance, *Biometrics* **35**, 757–763.

Schoener, T. W. 1973. Population growth regulated by intraspecific competition for energy or time: some simple representations, *Theoret. Population Biol.* **4**, 56–84.

Schoener, T. W. 1978. Effects of density-restricted food encounter on some single-level competition models, *Theoret. Population Biol.* **13**, 365–381.

Seradilla, J. 1979. "Population Growth and Fitness in *Drosophila*," Ph.D. Dissertation, University of California, Davis.

Smith, F. E. 1963. Population dynamics in *Daphnia magna* and a new model for population growth, *Ecology* **44**, 651–663.

Thomas, W. R., Pomerantz, M. J., and Gilpin, M. E. 1980. Chaos, asymmetric growth and group selection for dynamical stability. *Ecology* **61**, 1312–1320.

Tracey, M. L., and Ayala, F. J. 1974. Genetic load in natural populations: Is it compatible with the hypothesis that many polymorphisms are maintained by natural selection, *Genetics* **77**, 569–589.

Volterra, V. 1931. "Leçons sur la Théorie Mathématique de la Lutte pour la Vie," Gauthier–Villars, Paris.

Walls, R. C., and Weeks, D. L. 1969. A note on the variance of a predicted response in regression. *Amer. Statistician* **23**(3), 24–26.

Printed by the St. Catherine Press Ltd., Tempelhof 41, Bruges, Belgium