Res. Popul. Ecol. (1986) 28, 267–280.

ESTIMATION OF STAGE-SPECIFIC DEVELOPMENTAL TIMES AND SURVIVORSHIP FROM STAGE FREQUENCY DATA

S. M. SCHNEIDER[1] and H. FERRIS[2]

Department of Nematology, University of California, Riverside, California 92521, U.S.A.

INTRODUCTION

Life history studies usually include the determination or estimation of recruitment (actual number and/or distribution function), duration, and survivorship for each life stage. For single or discrete generation cases, many methods for calculating these values have been described (ASHFORD et al., 1970; BELLOWS and BIRLEY, 1981; BIRLEY, 1977; BIRLEY, 1979; HOGG and NORDHEIM, 1983; KIRITANI and NAKASUJI, 1967; KOBAYASHI, 1968; MANLY, 1974a; MANLY, 1974b; MANLY, 1976; MANLY, 1977; MILLS, 1981a; MILLS, 1981b; RICHARDS et al., 1960; SAWYER and HAYNES, 1984). In studies of above-ground, macroscopic, caged populations, the same individuals can be observed over time. In some cases, individuals can be observed repeatedly as they develop throughout their life cycle. Sampling bias is not introduced. For many soil organisms and microorganisms, destructive sampling is necessary. Population assessment techniques are not 100% efficient. Several individuals should be present in each sampling unit to ensure a reasonable probability of recovery. Population assessment efficiency varies with each life stage. Stage-specific assessment efficiencies can be estimated and used to adjust data. Data sets consisting of means of adjusted counts obtained through destructive sampling are more variable than those based on direct observations of individual development. Additional variation occurs in multiple generation data sets due to inherent genetic and microhabitat-induced differences in fecundity rate and length of the reproductive period. Estimation of recruitment, stage duration, or survivorship from such variable data which may not have generations sufficiently discrete to separate the same stage in successive generations, is diffiicult. This paper presents a basic age-structured, distributed delay population model and develops a parameter estimation algorithm. The algorithm is tested on a single generation data set of an insect pest (the simple case) (BELLOWS and BIRLEY, 1981) and applied to a plant parasitic nematode data set extended beyond the first generation (SCHNEIDER and FERRIS, 1987).

[1] Present address: USDA-ARS-SAA, Tobacco Research Laboratory, P.O. Box 1555, Oxford, North Carolina 27565, U.S.A.

[2] Present address: Division of Nematology, University of California, Davis, California 95616, U.S.A.

Population Model Structure

The general structure of the population model is

$$N_{i,t}=N_{i,t-1}+acN_{i-1,t-1}-bN_{i,t-1} \quad (1)$$

where $N_{i,t}$ is the number of individuals in stage i at time t, a is the proportion of individuals entering stage i from the previous stage, c is the proportion of individuals surviving the passage from stage i-1 to stage i, and b is the proportion of individuals maturing from stage i to stage $i+1$. In deterministic models in which all individuals of the same age move on to the next stage once the appropriate length of time (chronological or physiological) has passed, coefficients a and b are either 0 or 1. A zero indicates enough time has not passed and the probability of an individual moving to the next stage is zero. A one indicates maturity has been reached, and the probability of all individuals in that age cohort moving to the next stage is one.

In Leslie matrix format, the organism's life cycle is divided into substages with a length equal to the length of the shortest stage. The population is advanced at time steps equal to this substage length. At each time step, all individuals which survive a given substage are advanced to the next substage. The a and b coefficients are always 1. Equation 1 can be written:

$$N_{i,t}=N_{i,t-1}+1cN_{i-1,t-1}-1N_{i,t-1} \quad (2)$$

$$N_{i,t}=cN_{i-1,t-1} \quad (3)$$

The coefficient c corresponds to $(1-q)$ in life table notation (Leslie, 1945; Leslie, 1948). In a distributed delay model, movement of an individual from one stage to the next is not only based on the mean length of time in the stage, but also the variability among individuals associated with that mean. Some individuals mature faster than the mean, most at about the mean, and some slower than the mean (Ashford et al., 1970; Curry et al., 1978; Manetsch, 1976; Sharpe et al., 1977; Wagner et al., 1984). The a and b coefficients of equation 1 are calculated from a probability distribution based on the mean and standard deviation or are assigned a predetermined probability.

Survivorship can follow one of three general patterns; 1) high mortality at the beginning of the stage, 2) mortality evenly distributed across the stage, or 3) high mortality at the end of the stage (Pearl and Miner, 1935). A function describing the survivorship pattern characteristic of the organism being modeled can be chosen and implemented in equation 1. The form of the equation presented here reflects the third pattern, high mortality at the end of the stage. If a general function that can take on all three characteristic patterns is needed, then the Weibull function can be used (Hogg and Nordheim, 1983; Pinder et al., 1978; Wagner et al., 1984).

Developmental rates in ectothermic organisms are dependent on temperature and concentrations of developmental enzymes (Curry et al., 1978; Sharpe et al., 1977). Enzyme concentration can be assumed to be symmetrically distributed about some mean within a population of individuals. The symmetrical distribution of enzyme concen-

tration gives rise to symmetrically distributed developmental rates within the population. When developmental rates are converted to developmental time (*Time*$=1/$*rate*), the resulting distribution is asymmetric with a drawn out right-hand tail. This suggests that the choice of a probability function to govern maturation from one stage to the next should be an asymmetric, positively-skewed function. The function chosen for the model presented here is the ERLANG distribution. This function was chosen because it is positively skewed, it is completely defined by only two parameters, the mean and variance, it conserves flow, and it is relatively easily implemented on a computer (ASHFORD et al., 1970; MANETSCH, 1976). The equation for the ERLANG function is:

$$f(t)=(\bar{x}/k)^{-k}(t)^{k-1}\exp[-kt/\bar{x}]/(k-1)!$$

where $\bar{x}$ is the mean, s^2 is the variance, $k=\bar{x}^2/s^2$, and t is time (chronological or physiological). When $k=1$, this is the exponential distribution. As k approaches infinity, the ERLANG distribution approaches a normal distribution (MANETSCH, 1976).

The conceptual population model (Fig. 1) shows the progression of individuals through the life stages from egg to adult, with the probability of maturing to the next stage described by the ERLANG function. Each life stage is divided into k substages. A stability condition to insure flow is conserved requires that the time step (DT) used to calculate movement from one substage to the next for each stage be chosen such that $DT\leq\bar{x}/2k$. Since $\bar{x}/k$ is the length of time in each substage of a stage, this restriction insures that an individual may move only to the next substage during a single iteration. The parameters $\bar{x}$ and k will vary between stages, resulting in a different maximum allowable DT for each stage. The smallest of the allowable stage-specific time steps will meet the stability requirement for all other stages and is chosen as the time step to be used for the whole model. Mortality in this model is assumed to be distributed uniformly across the stage, and is implemented by multiplying the k^{th} root of total stage survivorship by the number of individuals leaving each substage. This could be changed to accommodate other mortality patterns.

Given a mean, variance, and survivorship for each stage, a fecundity rate for the adult stage, and the starting population age structure, the progress of individuals through the life cycle is updated at time intervals dictated by the time step. Updating occurs in reverse order, beginning with the oldest individuals and proceeding to the youngest, to prevent instantaneous graduation. Instantaneous graduation occurs when individuals mature from stage i to $i+1$ during the updating of stage i, and are then immediately

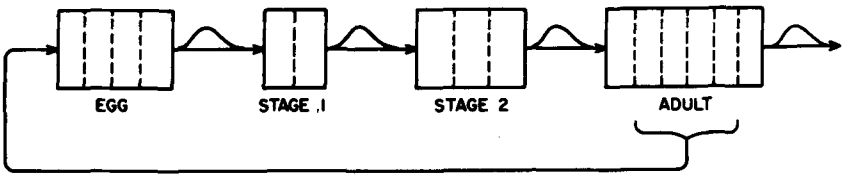

Fig. 1. Conceptual model of the progression of individuals through their life cycle. The broken lines indicate the k substages within a stage.

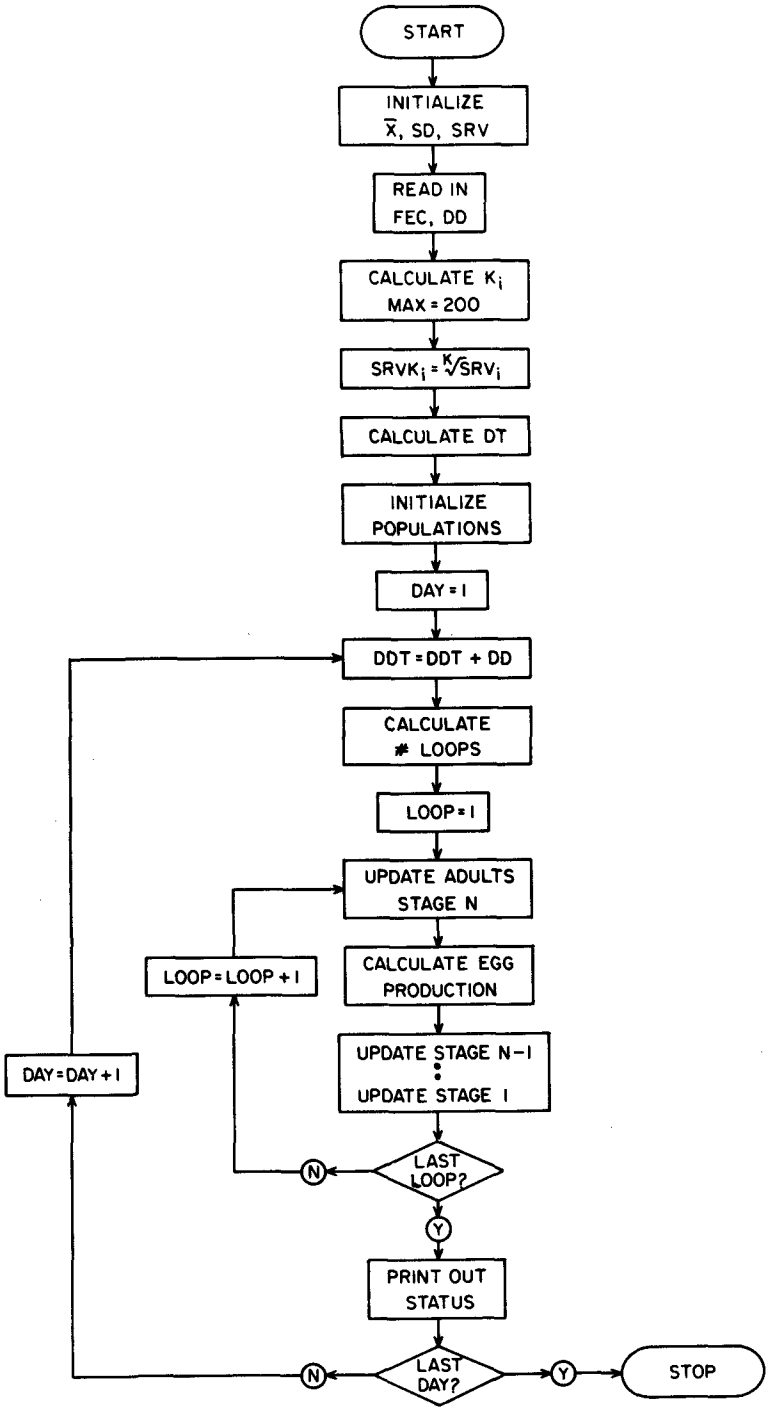

Fig. 2. Flowchart of an age-structured, distributed delay population model.

updated again during the updating of stage $i+1$. Adults are updated first as dictated by the mean and variance for adults and the time step. Next, egg production by the adults is calculated. All juvenile stages are updated based on their means and variances. Finally, the egg stage is updated, and the new eggs produced during this time interval are put into the first egg substage. As individuals pass from one substage to the next, they are multiplied by the sub-stage survivorship. A flowchart of this program is given (Fig. 2).

Estimation of Parameters

The general approach to parameter estimation is to generate predicted population values using a particular combination of parameters and compare the predicted to observed data. A weighted least squares (LSQ) is used to evaluate the goodness-of-fit

$$LS_iQ = \sum_{k=1}^{t} \sum_{j=1}^{n} (O_{j,k} - P_{j,k})^2 / O_{j,k} \tag{4}$$

where i represents a particular parameter combination, n is the number of stages, t is the number of time intervals, O is the observed population level, and P is the predicted population level. The parameter combination which minimizes the least squares value is chosen as the best estimate of the parameters

$$LSQ_{best} = \text{MIN}\ (LSQ_1,\ LSQ_2\ \ldots). \tag{5}$$

The algorithm is best suited to data sets in which the initial population is a uniform age cohort of first stage individuals, observations have been made at time intervals shorter than the expected length of the shortest stage, and the fecundity rate is known. Even with these conditions met, the number of parameters to be estimated can be a problem. If four values are tested for each of three parameters for four stages, the total number of combinations is 16,777,216 (4^{12}). If a time step of 1 degree-day (a physiological measure of time denoted DD) is chosen for a data set of 500 DD, this results in 8,388,608,000 iterations of the simulator. Even with a computer, this can be time consuming.

A stepwise procedure was developed to alleviate this problem (Fig. 3). If the data set contains more than one generation, a three-step process is used. The first generation data are used to estimate a mean developmental time for each stage, then the associated stage-specific variance, and survivorship. Finally, the full data set is used to refine all estimates.

In the initial parameter estimation phase, fecundity is set to zero. Any parameter values which have been determined experimentally are set. The stage-specific standard deviation is initialized at 0.5 times the mean, allowing coarse resolution in initial parameter estimation. A range of values is chosen to bracket the probable mean duration of the stage. Timing of the observed and predicted population peaks for the first stage are compared until a best estimate for the first stage mean is obtained based on the smallest least squares value. The value for the first stage mean is set, and the whole procedure is

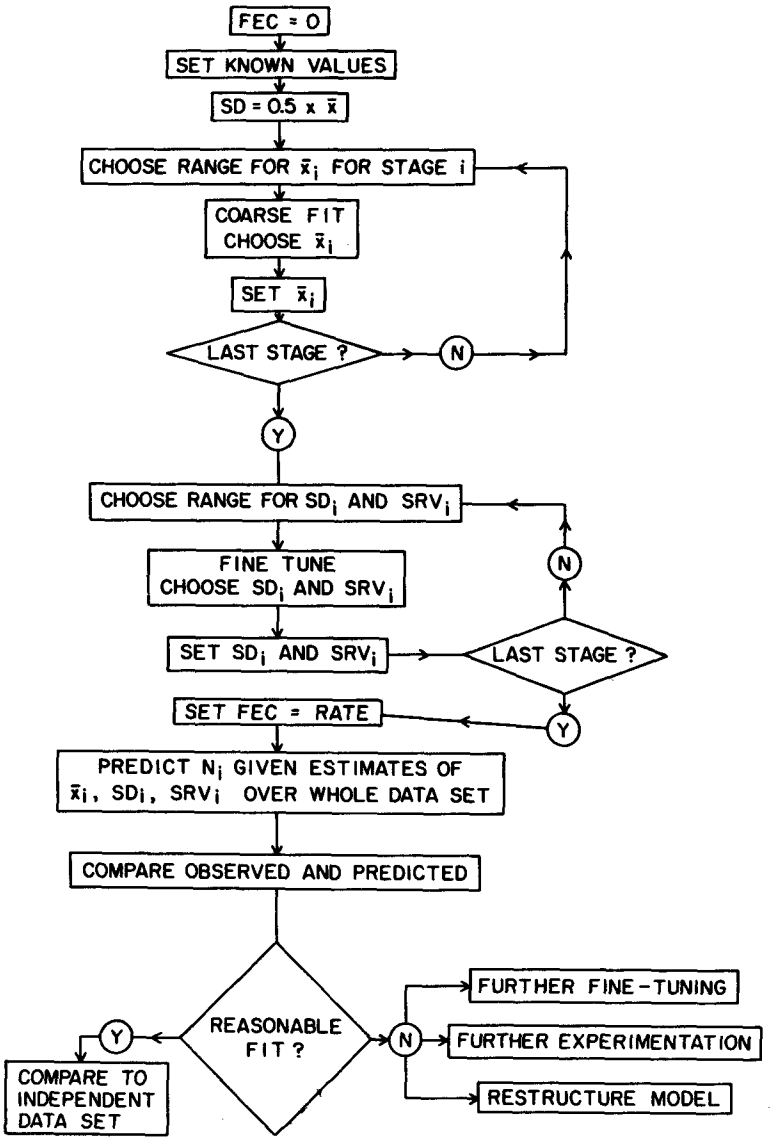

Fig. 3. Flowchart of a parameter estimation procedure.

repeated in a stepwise fashion until the mean duration of each stage has been estimated for all stages.

In the second step of parameter estimation, a range of values is chosen to breackt the probable standard deviation and survivorship of the first stage. With the best fit mean duration for the first stage, a predicted data set is generated for each standard deviation and survivorship combination. The parameter values minimizing the least squares value are selected. Fine resolution is obtained by narrowing the range of parameter values. Minor modification of the mean stage duration estimates may improve the fit. The procedure is repeated stepwise for all remaining stages.

In the third step, fecundity is set to its known value. Estimated parameter values from the first generation are used to simulate multiple generations. Observed and predicted values are compared as before. If the fit is not adequate, the parameter values are not verified, and further fine-tuning of the parameter estimates may be necessary. This might include testing values outside the ranges used previously. Further experimentation may be indicated if large variation in the data hinders the determination of best parameter estimates. In the case of large discrepancies between the observed and predicted values, a restructuring of the conceptual population model may be required. If the fit is satisfactory, the parameter values are verified, and the model and estimated parameters are ready for validation against an independent data set.

Application of the Parameter Estimation Procedure

A life table for the insect *Callosobruchus chinensis* (L) (Bellows and Birley, 1981) contains population counts for an egg stage, four juvenile stages, a pupal stage, and an adults stage over 26 days. Only one adult stage, the pre-emergent adults, was used for comparison. This is a single generation data set and no reproduction occurred. The first two steps of the parameter estimation procedure, coarse-fitting and fine-tuning using first generation data, were applied to this data set. The egg population was initialized as a uniform age cohort into the first egg substage. Developmental times and survivorship estimates by Bellows and Birley (1981) were compared with those obtained by the procedure described here.

The technique was also applied to an ectoparasitic plant nematode, *Paratrichodorus minor* (Colbran, 1956) Siddiqi, 1974 (Schneider and Ferris, 1987). The data set contained observations, corrected for extraction efficiency, for four juvenile stages and an adult stage over 378 degree-days using a basal threshold of 10C (DD_{10}). This data set extends beyond the first generation. Some parameters had been determined experimentally; the fecundity rate was 0.784 progeny/female/DD_{10} and the length of the egg stage was 53 ± 7.4 DD_{10}. All three steps of the parameter estimation procedure (coarse-fitting of the stage mean durations using first generation data, fine-tuning of the stage-specific standard deviations and survivorships from the first generation data, and extrapolation using the fecundity rate to the full data set) were applied to the data. The first stage juveniles were initialized as a uniform age cohort into the first substage. Comparisons between observed and predicted data were made for the first generation estimates, and, after fine tuning, for the whole data set.

Results

C. chinensis Data Set

Algorithm determined parameter values (Table 1) allowed population predictions close to observed data for the egg, first larval, and pupal stages (Fig. 4a, 4b, 4c). Pre-

Table 1. Estimates of the mean developmental time in days and survivorship for *Callosobrichus chinensis* using data from BELLOWS and BIRLEY (1981).

Stage	Length of Stage (days)		Survivorship (stage total)	
	BELLOWS and BIRLEY[a]	This paper[b]	BELLOWS and BIRLEY	This paper
Egg	3.16	4.10	1.00	0.98
1st larvae	2.13	2.62	1.00	0.93
2nd larvae	2.07	2.30	0.99	0.98
3rd larvae	2.06	1.95	1.00	1.00
4th larvae	4.43	4.95	1.00	1.00
Pupae	3.17	3.55	0.95	0.96
Adults in pea	2.11	2.60	1.00	1.00

[a] Using discrete GAUSSIAN distribution (BELLOWS and BIRLEY, 1981)
[b] Using an ERLANG probability distribution

dicted populations were slightly higher than observed for the second larval stage (Fig. 4c), and slightly lower than observed for the third and fourth stage larvae and the adults (Fig. 4d, 4e, 4g).

Estimates of developmental times calculated from the algorithm described here were generally higher than those estimates by BELLOWS and BIRLEY (1981) (Table 1). The duration of the stage was estimated to be 30%, 23%, 11%, 12%, 12% and 23% longer for the egg, first larval, second larval, fourth larval, pupal, and adult stages respectively. The third larval stage was estimated to be 5% shorter. While stage length was generally estimated to be longer, total stage survivorship was estimated to be lower in the first three life stages (Table 1).

P. minor Data Set

Estimates for stage mean duration, standard deviation, and survivorship following the coarse-fit and fine-tuning steps for the first generation (Table 2) provided minimum deviation between observed and predicted populations (Fig. 5). When parameter values were used to predict beyond the first generation, predicted second generation first stage juveniles entered the system too early. The model was restructured to include a pre-ovipositional adult period with a fixed length of 79 DD_{10} (SCHNEIDER and FERRIS, 1987).

Table 2. Estimates of Parameter Values for *P. minor*.

Stage	Mean (DD_{10})		S.D. (DD_{10})		Survivorship (total stage)	
	Original	Modified	Original	Modified	Original	Modified
Egg	53	53	7.4	7.4	1.00	1.00
J_1	22	14	6.7	4.2	1.00	0.90
J_2	44	50	4.4	7.5	1.00	0.87
J_3	47	45	4.7	6.7	1.00	0.90
J_4	14	8	1.4	3.2	1.00	0.90
Pre-Adult	—	79	—	23.7	—	1.00
Adult	78	100	39.0	70.0	1.00	0.50

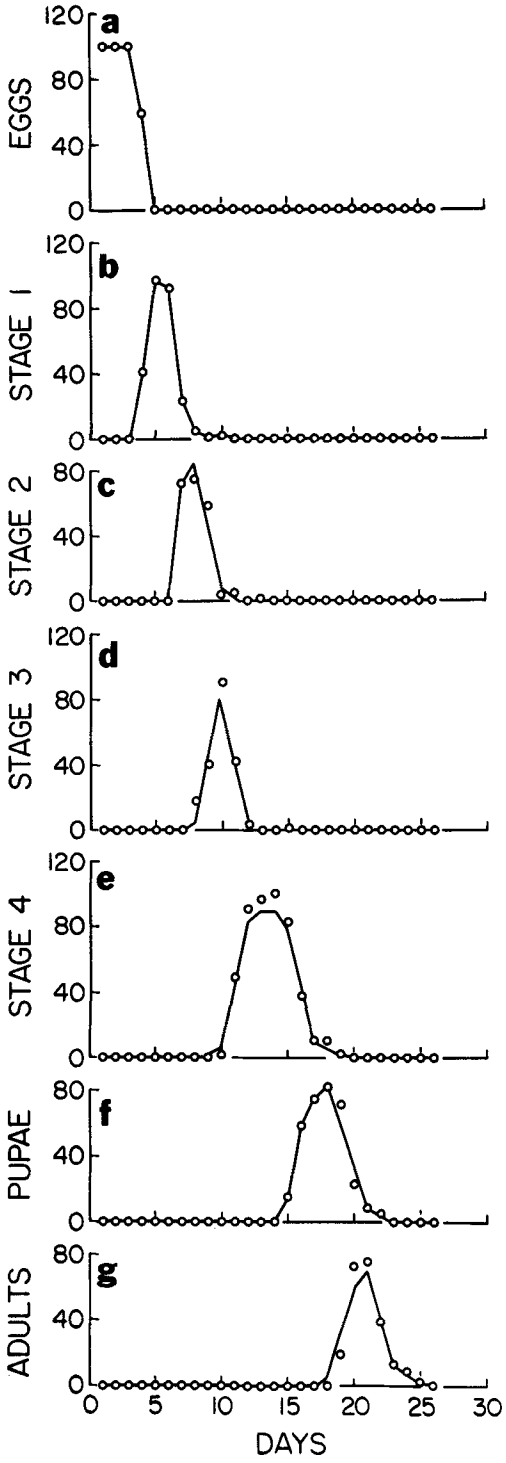

Fig. 4. Observed and predicted data for *Callosobruchus chinensis*. O are observed data points, — are predicted values.

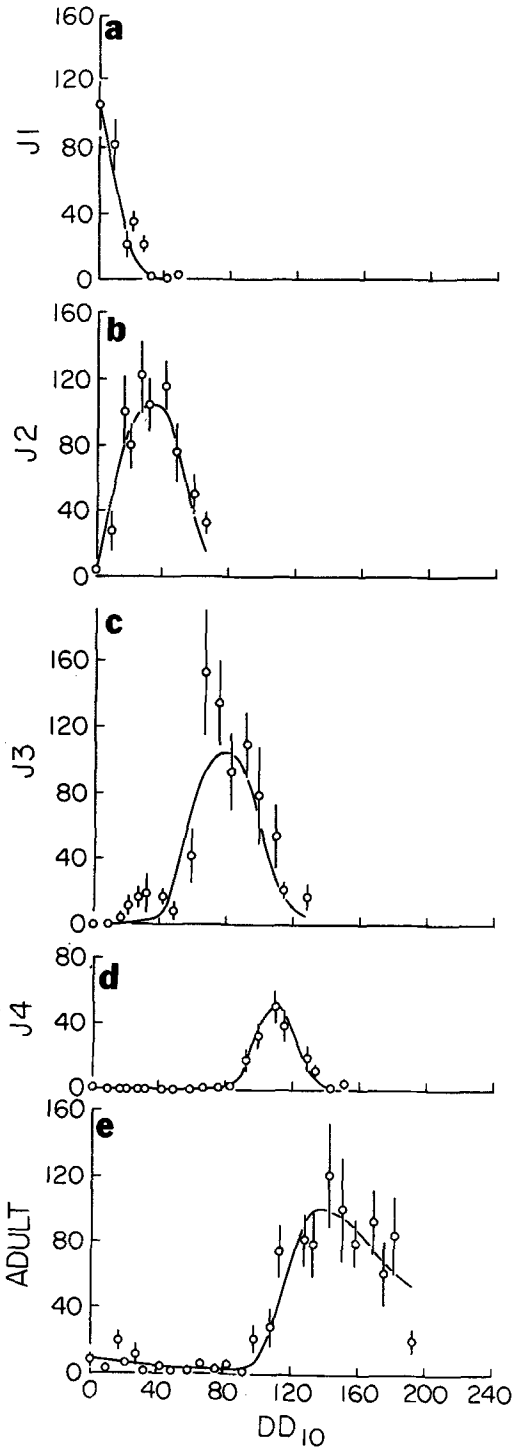

Fig. 5. Observed and predicted data for the first generation data from *Paratrichodorus minor*. O are observed data points, mean of ten replications with standard error bars, — are predicted values.

Values of the stage mean durations, standard deviations, and survivorships, were further refined using the whole data set, ultimately yielding modified values (Table 2). Predicted first stage juvenile populations are higher than observed from 220–320 DD_{10} and

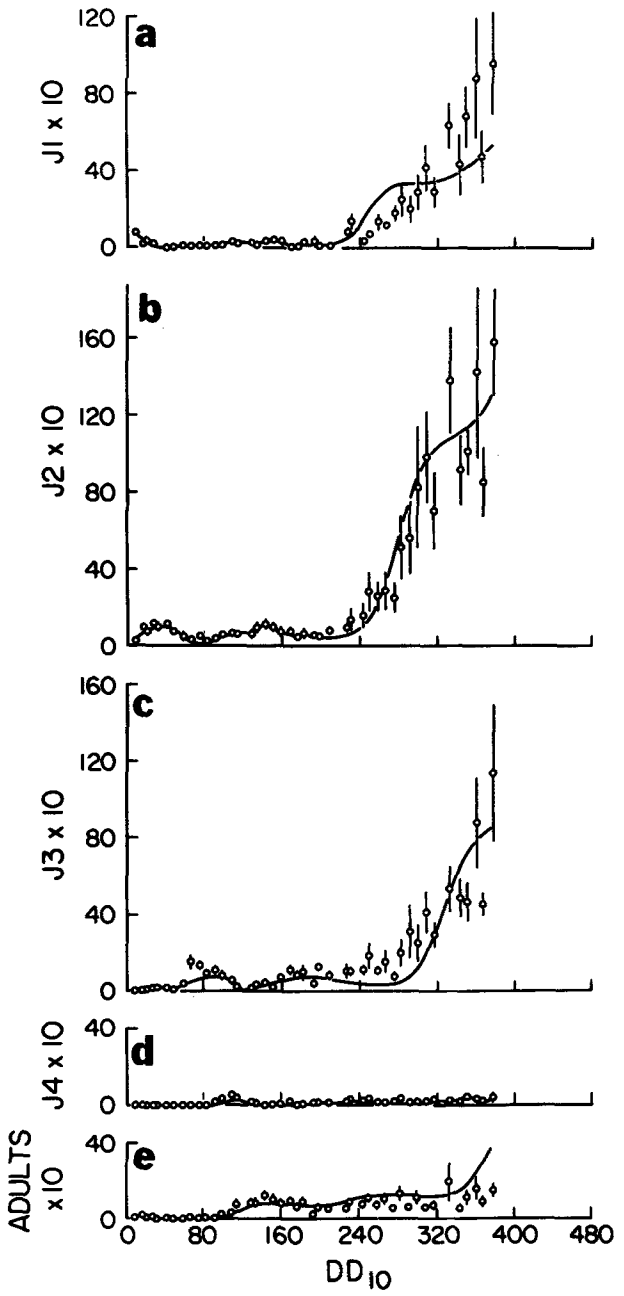

Fig. 6. Observed and predicted data for the full data set of *Paratrichodorus minor*. O are observed data points, mean of ten replications with standard error bars, — are predicted values.

lower than the observed from 320–378 DD_{10} (Fig. 6a). Predicted second stage juvenile populations are not systematically higher or lower than observed over any time interval (Fig. 6b). Predicted populations of third stage juveniles are lower than observed from 120–320 DD_{10} (Fig. 6c). Predicted populations of fourth stage juveniles are higher than observed from 330–378 DD_{10} (Fig. 6a). Predicted adult populations are lower than observed from 100–190 DD_{10} and higher than observed from 190–378 DD_{10} (Fig. 6e).

Discussion

The method of life table parameter estimation proposed in this paper worked well in determining parameter values for a single generation insect data set. The values obtained were not the same as those estimated previously (Bellows and Birley, 1981) but were in general agreement. The differences which occurred are probably due to the choice of different probability functions in the population model. Both functions can be used to represent the observed data satisfactorily.

There was more variation between observed and predicted values in the multiple generation nematode data set. This is to be expected since cumulative error is a function of the number of stages and the length of time simulated. Variation in extraction efficiencies can lead to errors in the estimation of population numbers (Schneider and Ferris, 1987), and hence estimation of life table parameters. Inaccuracies in the determination of any preset value, such as fecundity, will also affect the parameter estimates.

The initial age distribution of the population is another possible source of error. A population can be considered as a uniform age cohort in the first substage of a stage, for example, right after a synchronous egg hatch. The initial population might also be a uniform age cohort in the last substage of a stage, as before a synchronous egg hatch. The distribution of individuals across all substages according to a predetermined age distribution would be another option. The most valid choice may vary for different data sets.

A disadvantage of this parameter estimation procedure is that it does not estimate both fecundity and stage survivorship. These parameters are interrelated and overestimation of fecundity can be offset by underestimating survivorship. Underestimating fecundity is compensated for by overestimating survivorship until survivorship is 100% for all stages. However, if the stage-specific survivorships can be determined experimentally, these can be set as fixed values and the fecundity rate estimated by the parameter estimation procedure.

The parameter estimation algorithm presented here uses a population model with an Erlang probability distribution to choose values for stage-specific mean development times, the standard deviations associated with the development times, and stage specific survivorships. The algorithm requires a data set containing stage frequencies over time, ranges for the values to be estimated, and fecundity rate, if the data set extends beyond the first generation. If a large number of parameters are to be estimated, a single optimal

solution may not be found. The mathematically optimal solution may not be the most biologically reasonable choice. In determining the acceptance or rejection of any parameter based solely on mathematical best fit, biological common sense must be used.

Summary

An algorithm for estimating mean stage durations, the standard deviations associated with the means, and stage-specific survivorships from stage frequency data is presented. The algorithm is based on an age-structured, distributed delay simulation model which uses Erlang distributions to determine the probability of maturity for individuals in each stage. If the data set extends beyond the first generation, the algorithm requires a fecundity rate, as well as stage frequencies, as input. Goodness-of-fit was measured using a weighted least squares calculation summed over all observed stages and all sampling dates.

References

Ashford, J. R., K. L. Q. Read and G. G. Vickers (1970) A system of stochastic models applicable to studies of animal population dynamics. *J. Anim. Ecol.* **39**: 29–50.

Bellows, T. S. Jr. and M. H. Birley (1981) Estimating developmental and mortality rates and stage recruitment from insect stage frequency data. *Res. Popul. Ecol.* **23**: 232–244.

Birley, M. (1977) The estimation of insect density and instar survivorship function from census data. *J. Anim. Ecol.* **46**: 497–501.

Birley, M. (1979) The estimation and simulation of variable developmental period, with application to the mosquito *Aedes aegypit* (L.). *Res. Popul. Ecol.* **21**: 68–80.

Curry, G. L., R. M. Feldman and P. J. H. Sharpe (1978) Foundations of stochastic development. *J. Theor. Biol.* **74**: 397–410.

Hogg, D. M. and E. V. Nordheim (1983) Age-specific survivorship analysis of *Heliothis* spp. populations on cotton. *Res. Popul. Ecol.* **25**: 280–297.

Kiritani, K. and F. Nakasuji (1967) Estimation of the stage-specific survival rate in the insect population with overlapping stages. *Res. Popul. Ecol.* **9**: 143–153.

Kobayashi, S. (1968) Estimation of the individual number entering each development stage in an insect population. *Res. Popul. Ecol.* **10**: 40–44.

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.

Manetsch, T. J. (1976) Time-varying distributed delays and their use in aggregative models of large systems. *IEEE Trans. Systems, Man, and Cybernetics.* Vol. SMC **6**: 547–553.

Manly, B. F. J. (1974a) Estimation of stage-specific survival rates and other parameters for insect populations developing through several stages. *Oecologia* **15**: 277–285.

Manly, B. F. J. (1974b) A comparison of methods for the analysis of insect stage-frequency data. *Oecologia* **17**: 335–348.

Manjy, B. F. J. (1976) Extensions to Kiritani and Nakasuji's method for analyzing insect stage-frequency

data. *Res. Popul. Ecol.* **17**: 191–199.

MANLY, B. F. J. (1977) A further note on KIRITANI and NAKASUJI's model for stage-frequency data including comments on the use of TUKEY's jackknife technique for estimating variance. *Res. Popul. Ecol.* **18**: 177–186.

MILLS, N. J. (1981a) The estimation of mean duration from stage frequency data. *Oecologia* **51**: 206–211.

MILLS, N. J. (1981b) The estimation of recruitment from stage frequency data. *Oecologia* **51**: 212–216.

PEARL, R. and J. R. MINER (1935) Experimental studies on the duration of life. XIV. The comparative mortality of certain lower organisms. *Quart. Rev. Biol.* **10**: 60–79.

PINDER, J. E., III, J. G. WIENER and M. H. SMITH (1978) The Weibull distribution: a new method of summarizing survivorship data. *Ecology* **59**: 175–179.

RICHARDS, O. W., N. WALOFF and J. P. SPRADBERY (1960) The measurement of mortality in an insect population in which recruitment and mortality widely overlap. *Oikos* **11**: 306–310.

SAWYER, A. J. and D. L. HAYNES (1984) On the nature of errors involved in estimating stage-specific survival rates by Southwood's method for a population with overlapping stages. *Res. Popul. Ecol.* **26**: 331–351.

SCHNEIDER, S. M. and H. FERRIS (1987) Stage-specific population development and fecundity of *Paratrichodorus minor*. *J. Nematol.* **19**: (in press).

SHARPE, P. J. H., G. L. CURRY, D. W. DEMICHELE and C. L. COLE (1977) Distribution model of organism development times. *J. Theor. Biol.* **66**: 21–38.

WAGNER, T. L., H. WU, P. J. H., SHARPE and R. N. COULSON (1984) Modeling distributions of insect development time: A literature review and application of the Weibull function. *Ann. Entomol. Soc. Amer.* **77**: 475–487.

発育ステージ別の頻度データからステージごとの成長時間と生存率の推定

S. M. SCHNEIDER · H. FERRIS

発育ステージ別の頻度データから，平均ステージ期間とその標準偏差およびステージ別生存率を推定する1つの計算法を示した．この方法は，各ステージにおける個体の成長確率がERLANG分布で与えられるものとして，遅れをともなう分布をもつ令構成のシミュレーション・モデルに基づいている．1世代を越えるデータの場合，ステージ別頻度のほかに繁殖率を入力する必要がある．観察されたすべてのステージとすべてのサンプリング日について合計した加重最小2乗値で適合度を測った．