

The Estimation of Mean Duration from Stage Frequency Data

N.J. Mills

Hope Department of Entomology, University Museum, Oxford, England

Summary. A simple method of estimating duration from stage frequency data is derived. A simulation model of the passage of individuals through a particular stage in the life-cycle is presented, together with results from the model on the influence of recruitment, development and mortality on the parameters used in the estimation of stage duration. The application of the method to field data is described and a test example, using simulated data, is given.

Introduction

The durations of the various life-cycle stages of an insect collectively determine the generation time, an important component of population growth and thus the analysis of population dynamics. The influence of temperature on development rates has long been recognised and this has led to the production of temperature-rate functions for development from data on insects held at constant temperatures. In this way, stage durations are predicted for field populations from mean daily temperature records (e.g. Logan et al. 1976) or from a physiological accumulated temperature above a threshold (e.g. Davidson 1944; Campbell et al. 1974; Stinner et al. 1974).

There are several reasons to suspect the applicability of such temperature-rate functions of insect development to field populations. The first is that development rates tend to be more rapid under fluctuating temperature conditions, characteristic of the field (e.g. Hagstrum and Hagstrum 1970; Siddiqui and Barlow 1973). Secondly, the actual temperatures experienced by insects in the field may not be comparable to ambient air temperature due to the absorption of sunlight (e.g. Digby 1955; Woodburn et al. 1978).

Rates of development may also change from one location to another due to variations in temperature thresholds or thermal requirements for development in local populations (e.g. Morris 1971) or due to variations in other climatic factors. Finally, development rates vary with the quantity and quality of the food source (e.g. Dowell and Steinberg 1979; Mukerji and Guppy 1970).

For these reasons, temperature-rate functions of insect development tend to be specific to the conditions under which they were derived. It is therefore of considerable advantage in the analysis of population dynamics to estimate stage durations under natural conditions in the field. Despite this, relatively little effort has gone into the estimation of stage duration from census data (Manly 1977; Birley 1979), compared with that for the estimation of stage mortality (see Southwood 1978 for review). In the present paper, a simple method is derived for the estimation of stage duration from the frequency data obtained for each stage from a series of successive samples.

Derivation of Method

Consider a stage in the life-cycle of an insect population which suffers no mortality, has a constant duration of 4 days and is censused at daily intervals. An individual recruited to the stage at any time between the 1st and 2nd samples becomes part of the cohort recruited during day 2. This cohort is then present in the population at sample times 2, 3, 4 and 5 and is recruited to the succeeding stage during day 6. Thus each individual recruited to the stage is available for sampling on 4 successive occasions.

In Table 1 a simple example of the passage of daily cohorts through a stage, illustrates the derivation of the relationships to be used in the analysis of stage frequency data. Ten individuals are recruited to the stage over a 3 day period, giving a 'mean' time of recruitment of 3.1. The total number of individuals present in the stage each day (summed over the cohorts of recruitment) forms the sample curve, which has a 'mean' time of 4.6. The ten individuals are then recruited to the succeeding stage at a 'mean' time of 7.1.

The 'mean' time of the sample curve (MS) is therefore related to the 'mean' time of recruitment (MR) to the stage by:

$MS = MR + 0.5d - 0.5$	
$4.6 = 3.1 + 2 - 0.5$	(1)

Table 1. A simple example of the passage of daily cohorts of recruits (C) through a stage in the life-cycle of an insect population, with no mortality and a constant 4 day duration

Present Address: Department of Zoology, South Parks Road, Oxford 0Xt 3PS, England

while the 'mean' times of recruitment to two consecutive stages are given by:

$$
MR_{i+1} = MR_i + d_i
$$

7.1 = 3.1 + 4 (2)

where d is the stage duration. The 'mean' time of the sample curve is therefore shifted along the time axis by $0.5d - 0.5$ from the 'mean' time of recruitment and the succeeding 'mean' time of recruitment is shifted along by d . These relations hold for a variety of recruitment distributions, hence the use of the term 'mean'. The median time produces relations of the same form only for continuous distributions of recruitment.

For any two consecutive stages of the life-cycle (*i* and $i+1$) the appropriate equations for MS are:

$$
MS_i = MR_i + 0.5d_i - 0.5
$$
 (3)

and

$$
MS_{i+1} = MR_{i+1} + 0.5d_{i+1} - 0.5
$$

= MR_i+d_i+0.5d_{i+1}-0.5 (4)

since $MR_{i+1} = MR_i + d_i$ [Eq. (2)]. Then from (4) – (3)

$$
2(MS_{i+1} - MS_i) = d_i + d_{i+1}
$$
\n(5)

From Eq. (5) it can be seen that twice the difference in 'mean' time of the sample curve for two consecutive stages represents the sum of their stage durations. This therefore provides a useful basis for the determination of duration from stage frequency data.

However, this argument is based on the assumptions of no mortality and constant stage duration. In field populations these two assumptions will not hold, since insects characteristically show considerable mortality between stages (e.g. Podoler and Rogers 1975) and skewed distributions of development times (e.g. Sharpe et al. 1977). The influence of mortality and development distributions on the basic relationships expressed by Eq. (1) and (2) were investigated by the use of a computer simulation model.

Simulation **Model**

A discrete Fortran model was used to provide data on the numbers present in a particular stage in relation to recruitment, age-specific mortality and duration. The model is presented in Fig. 1 as a flow chart, indicating the various steps involved. Recruitment and mortality are up-dated each day at the time when the population is censused for the numbers present in the stage. A discrete model allows for the incorporation of a wide variety of recruitment distributions and survivorship functions, however, it must approximate very closely the more continuous nature of these processes under natural field conditions.

Recruitment in field populations may occur with peaks at particular times of the day, or may occur randomly throughout the day. In the model, each day's recruitment is incorporated into the population at the time when it is censused. There is therefore the problem of whether the recruits should have suffered any mortality at that time, when they are first censused. Assuming recruitment to occur randomly throughout the day leading up to the first census, then some mortality would certainly have occurred. This situation is well approximated by allowing each day's recruitment to suffer a level of mortality corresponding to an age of half a day. In Fig. 2 an age-specific survivorship function is shown, indicating the derivation of the values used for up-dating mortality at each census time in relation to the age of each day's cohort of recruitment.

Fig. 1. A flow chart of the discrete Fortran model used for the simulation of stage frequency data. $E(I)$ number recruited on day I ; $S F(L)$ survival rate at age L; *DF(L)* proportion remaining in stage at age *L;* $X(I, J)$ number recruited on day *I* present in stage on day *J*; $T(J)$ number present in stage on day J ; A area of sample curve; R total number recruited; *MS* 'mean' time of sample curve; *MR* 'mean' time of recruitment; D maximum age of stage; C last day of recruitment

Fig. 2. A diagram illustrating the derivation of the values used for the survivorship function in relation to age

Simulation Results

The ~Mean' Time of the Sample Curve

The shape of the recruitment distribution and the total number recruited to a stage do not influence the 'mean' time of the sample curve. However, the 'mean' time is influenced by the survivorship and development functions. Equation (1) can be modified to allow for these influences by:

$$
\frac{\text{MS}_i}{\text{Fa}_i} = \text{MR}_i + 0.5\,\overline{d}_i - 0.5\tag{6}
$$

where Fa_i is a correcting factor for the shift of the 'mean' time of the sample curve from the 'mean' time of recruitment and \bar{d}_i is the mean stage duration.

The effect of the survivorship function can be seen from the values of Fa in Fig. 3a for a stage of 10 days duration. The curves represent three simplified age-specific survivorship functions as shown in the inset graph. The effect of the development distribution was investigated using a four day range of stage duration giving a positively skewed distribution of the proportion of individuals changing stages on each of the four days. In Fig. 3b the value of Fa in the absence of mortality is given in relation to the mean stage duration for the two development distributions shown inset.

These graphs indicate that the correction term Fa decreases from unity as mortality increases and as the survivorship pattern changes to concentrate mortality towards the start of the stage. This reflects the decreased representation of non-survivors in the sample curve, with its consequent fall in the 'mean' time of the sample curve. In contrast to this, at shorter stage durations the positive skew of developmental distribution leads to values of Fa greater than unity. Individuals with greater than mean duration represent an increasing proportion of the sample curve as mean duration decreases, resulting in the 'mean' time of the sample curve being drawn further along the time axis.

The influence of both survivorship and development functions on the value of Fa can be combined in a multiple regression analysis. For this analysis the parameters used were mean duration (\overline{d}) and mean mortality rate (1- \overline{s}) since their curvilinear relationships with Fa are opposing and tend to cancel out when combined. Results were obtained from 100 stage simulations with varying survivorship and development functions, yielding the following regression equation:

$$
Fa = (1.0449 \pm 0.0023) - (0.0045 \pm 0.0002) \vec{d}
$$

- (0.0996 \pm 0.0027)(1 - \vec{s})

$$
r^2 = 0.942 \quad F = 791.3 \quad P < 0.001
$$
 (7)

This equation indicates that 94% of the variance of Fa can be accounted for by the mean duration and mortality rate of the stage.

The "Mean' Time of the Subsequent Recruitment Distribution

Again, as in the previous section, the recruitment distribution and total number recruited have no effect on the relationship of the 'mean' time of recruitment to two consecutive stages. There is a slight effect of the survivorship and development distributions on the relationship, which can be modified as:

$$
\frac{\text{MR}_{i+1}}{\text{Fb}} = \text{MR}_i + \vec{d}_i \tag{8}
$$

where Fb is a correction factor and \overline{d}_i is the mean duration of the i'th stage.

Fig. 3a, b. The influence of (a) the survivorship functions $1-3$ shown inset and (b) the development functions shown inset, on the value of the correction term Fa

There is no obvious trend in the values of Fb with mortality rate or mean duration and for a series of 50 stage simulations, under varying conditions of survivorship and development, a mean value of $0.998 + 0.003$ (S.D.) was obtained.

Application of the Method

The results from the simulation model suggest that Eq. (6) and (8) can be used with reasonable confidence for the estimation of stage duration from frequency data. In fact, the deviation of the correction factor Fb from unity in Eq. (8) is so small as to be negligable. Thus for two consecutive stages, twice the difference in corrected 'mean' time of the sample curves is equal to the sum of the two stage durations :

$$
2\left(\frac{\text{MS}_{i+1}}{\text{Fa}_{i+1}} - \frac{\text{MS}_i}{\text{Fa}_i}\right) = \overline{d}_i + \overline{d}_{i+1}
$$
\n
$$
(9)
$$

In order to use Eq. (9) values must be obtained for the correction term Fa for each stage. Since this is dependent on mean duration and mean mortality rate, which are unknown, direct estimates cannot be obtained. However, an intermediate value for Fa can be used for all stages, without introducing unacceptable error. From the survey of life-table data by Podoler and Rogers (1975), the survival rate for a single insect instar varies from 0.96 to 0.17. By substituing the mean of this range into Eq. (7), Fa is dependent on stage duration, as in Fig. 4. The mean duration of a generalised life-cycle stage can be estimated from :

$$
\bar{d} = \frac{\text{MS}_q - \text{MS}_1}{q - 1} \tag{10}
$$

Fig. 4. The dependence of Fa on mean duration $(----)$ for a stage with average survival rate, and the percentage error $(-)$ introduced by the assumption that $Fa = 1$

where there are q stages. An appropriate value of Fa can then be read from the graph in Fig. 4.

For smaller mean durations, Fa differs little from unity and the error introduced by such an assumption is small. However, the percentage error in the estimation of stage duration in this way increases with mean duration as shown in Fig. 4. Therefore, the use of an appropriate value of Fa is advantageous at all levels of stage duration though of greater significance as mean duration increases.

The use of a single value of Fa reduces Eq (9) to:

$$
\frac{2}{\text{Fa}}(\text{MS}_{i+1} - \text{MS}_i) = \bar{d}_i + \bar{d}_{i+1}
$$
\n(11)

From this equation a series of estimates of the sum of consecutive pairs of stage durations can be obtained. Then, provided that any one of the stage durations, or the ratio of any one pair of stage durations, can be independently estimated, the individual stage durations are found by difference. For many insect populations the sessile, non-feeding stages of the life-cycle (egg and pupa) may be most easily marked in the field, and their stage durations estimated directly. However, in some cases this may not be feasible and the duration of these stages, which are subject to less variability, could be independently estimated from temperature-rate functions. Alternatively, the ratio of the pre-pupal and pupal stage durations may be constant for some holometabolous insects, as in the coccinellid *Adalia bipunctata* (L.) (Mills 1979).

A Test Example of the Method

In order to assess the accuracy of this method of analysis, computer simulated population data were used. The model in Fig. 1 was adapted to provide data for a series of consecutive stages, requiring the input of a recruitment distribution for the first stage and survivorship and development data for all stages of the life-cycle.

Two separate population simulations were run, each with six stages to the life-cycle. The initial recruitment distributions

Table 2. The numbers present in each stage at each census and the 'mean' time of the sample curves (MS) for population l

Day	Census	Stage					
		$\mathbf{1}$	$\overline{2}$	3	$\overline{4}$	5	6
1	1	0.074					
	\overline{c}	0.207					
$\begin{array}{c} 3 \\ 5 \\ 7 \end{array}$	$\overline{\mathbf{3}}$	0.388					
	4	0.571	0.021				
9	5	0.470	0.067				
11	6	0.316	0.119	0.009			
13	$\overline{7}$	0.159	0.154	0.031	0.001		
15	8	0.077	0.109	0.056	0.013		
17	9	0.033	0.052	0.063	0.040		
19	10	0.010	0.027	0.044	0.063	0.005	
21	11	0.001	0.014	0.018	0.072	0.018	
23	12		0.004	0.010	0.053	0.040	
25	13			0.005	0.028	0.062	
27	14			0.001	0.014	0.068	0.002
29	15				0.007	0.059	0.010
31	16				0.002	0.042	0.022
33	17					0.023	0.034
35	18					0.011	0.041
37	19					0.005	0.041
39	20					0.002	0.037
41	21						0.031
43	22						0.022
45	23						0.013
47	24						0.006
49	25						0.003
51	26						0.001
	MS ^a	7.950	13.190	16.916	21.020	27,460	37.357

 $n_{ij}=N_{0}$, present in stage i at j'th census; $t_{ij}=D_{0}$ of j'th census; q =No. of census times

$$
M S = \frac{\sum_{j=1}^{q} n_{ij} t_j}{\sum_{j=1}^{q} n_{ij}}
$$

were irregular and multimodal over a period of 15 days; while survivorship patterns concentrated mortality towards both ends of the stage, and development distributions were positively skewed over a 4 day period. The populations were censused every 2 days to determine the number present in each stage, as shown in Table 2 for population 1. The resulting 'mean' time of each stage sample curve and recruitment distribution are given, together with appropriate input conditions in Table 3.

From the 'mean' times of the sample curves in Table 3, values of Fa for populations 1 and 2 were determined as 0.976 and 0.979 respectively, using Eq. (10). Estimates of the combined duration of consecutive pairs of stages, calculated using Eq. (11), are given in the first half of Table 4. For population 1 it was assumed that the duration of the first (egg) stage was known, from which the individual stage durations could then be obtained by difference (second half of Table 4). For population 2 it was assumed that the duration of stage 5 is a constant proportion (0.62) of the duration of stage 6. From this assumption \bar{d}_6 = 16.78/1.62 and then the individual stage durations can again be obtained by difference. The resulting estimates of the stage durations are in close agreement with the actual input values, as shown in Table 4, giving low values for the mean squared differences.

Stage	Input R	\tilde{s}	ā	Results МR	MS	
Population 1						
1 2 3 4 5 6	1.000 0.473 0.186 0.127 0.010 0.059	0.473 0.393 0.687 0.794 0.586 0.487	6.40 4.08 3.10 5.15 8 2 6 12.35	5.602 11.919 15.974 19.030 24.158 32.377	7.950 13.190 16.916 21.020 27.460 37.357	
Population 2						
1 $\overline{2}$ 3 $\overline{4}$ 5 6	1.000 0.636 0.245 0.195 0.132 0.098	0.636 0.386 0.796 0.675 0.746 0.802	6.15 4.20 3.01 4.30 6.36 10.26	8.368 14.486 18.582 21.574 25.835 32.175	10.702 15.918 19.657 23.232 28.481 36.696	

Table 3. The input conditions and results from the simulated population examples

Table 4, The combined duration of consecutive pairs of stages and the individual stage durations, estimated from the 'mean' time of the sample curves

Stage		Population 1	Population 2		
$1 + 2$		10.74	10.66		
$2 + 3$	7.64		7.64		
$3 + 4$	8.41		7.30		
$4 + 5$	13.20		10.72		
$5 + 6$	20.28		16.78		
	Actual	Estimated	Actual	Estimated	
1	6.40	6.40	6.15	6.02	
2	4.08	4.34	4.20	4.64	
3	3.10	3.30	3.01	3.00	
4	5.15	5.11	4.30	4.30	
5	8.26	8.09	6.36	6.42	
6	12.35	12.19	10.26	10.36	
$(Act - Est)^2$ n	0.027		0.037		

Discussion

This method of estimating stage duration from census frequency data has the advantages of simplicity (i.e. does not involve computation) and a requirement for the minimum of additional information (i.e. duration of any single or ratio of any one pair of stages). It can be applied to any insect or other invertebrate population which has generations sufficiently discrete to separate the same stage in successive generations and allows for considerable overlap of consecutive stages. The method requires that the census interval be sufficiently short to provide several frequency observations for each stage, the interval being dependent on the duration of recruitment and development for the stages involved. Differential efficiency of sampling of the various stages is unlikely to introduce errors, although the consistency of sampling at each census may determine the accuracy of the method.

Manly's method of analysis (Manly 1977) suggests that for a population of *n* stages, the duration of the *i*'th stage (d_i) can be estimated from:

 $d_i = B_{i+1} - B_i$

where B_i , is the 'mean' time of the sample curve for all stages i to *n* combined, and B_{i+1} the 'mean' time for stages $i+1$ to n combined. From Eq. (1), representing a simplified situation:

$$
B_i = MR_i + \left(0.5\sum_{i=1}^{n} d_i\right) - 0.5
$$

and

$$
B_{i+1} = MR_i + d_i + \left(0.5\sum_{i+1}^{n} d_i\right) - 0.5
$$

since

$$
MR_{i+1} = MR_i + d_i
$$

The difference

$$
B_{i+1} - B_i = d_i + \left(0.5 \sum_{i+1}^{n} d_i\right) - \left(0.5 \sum_{i}^{n} d_i\right) = 0.5 d_i
$$

If the correction factor *Fa* is taken into account the difference:

$$
B_{i+1} - B_i = 0.5 d_i + (Fa_{i+1} - Fa_i) \left(MR_i + 0.5 \sum_{i=1}^{n} d_i - 0.5\right)
$$

which in practice produces estimates which vary widely from the actual value of d_i . The value of the correction factor Fa decreases progressively as more stages are combined, increasing the difference between consecutive Fa values (i.e. $Fa_i < Fa_{i+1}$, and $Fa_{i+1} - Fa_i > Fa_{i+3} - Fa_{i+2}$).

The method of analysis suggested by Birley (1979) uses a rather different principle in that it provides an estimate of the development distribution for a stage by non-linear least squares regression of the recruitment distribution to two consecutive stages. While an estimate of the mean and distribution of stage duration is an important asset, the method has restricted application due to its requirement for the recruitment distribution to a stage. This must be extremely difficult, if not impossible, to obtain from the majority of field populations.

Therefore, the method of analysis described in the present paper is recommended for use by the field ecologist. It is preferable to that of Manly (1977) due to the unacceptable errors incorporated into the assumptions of that method. It also has advantages over the method of Birley (1979) in that the use of a computer is not needed, and the minimal additional information required is easily obtained.

Acknowledgements. I am grateful to the Ernest Cook Trust for financial support, and to Drs. D.J. Rogers and G.R.W. Wint and to Mr. C. Dye for useful comments on the manuscript.

References

- Birley MH (1979) The estimation and simulation of variable developmental period, with application to the mosquito *Aedes aegypti* (L.). Res Popul Ecol 21:68-80
- Campbell A, Frazer BD, Gilbert NE, Gutierrez AP, Mackauer M (1974) Temperature requirements of some aphids and their parasites. J Appl Ecol 11:431-438
- Davidson J (1944) On the relationship between temperature and rate of development of insects at constant temperatures. J Anim Ecol 13:26-38
- Digby PSB (1955) Factors affecting the temperature excess of insects in sunshine. J Exp Biol 32:279-298
- Dowell RV, Steinberg B (1979) Development and survival of immature citrus blackfly (Homoptera: Aleyrodidae) on 23 plant species. Ann Ent Soc Am 72:721-724
- Hagstrum DW, Hagstrum WR (1970) A simple device for producing fluctuating temperatures, with an evaluation of the ecological significance of fluctuating temperatures. Ann Ent Soc Am 63:1385-1389
- Logan JA, Wokind DJ, Hoyt SC, Tanigoshi LK (1976) An analytical model for description of temperature dependent rate phenomena in arthropods. Env Ent 5 : 1133-1140
- Manly BFJ (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 variances. Res Popul Ecol 18 : 177-186
- Mills NJ (1979) *Adalia bipunctata* (L.) as a generalist predator of aphids. Unpublished Ph.D. Thesis, University of East Anglia
- Morris RF (1971) Observed and simulated changes in genetic quality in natural populations of *Hyphantria cunea.* Can Ent 103:893-906
- Mukerji MK, Guppy JC (1970) A quantitative study of food consump-

tion and growth in *Pseudaletia unipunctata* (Lepidoptera: Noctuidae). Can Ent 102:1179-1188

- Podoler H, Rogers D (1975) A new method for the identification of key factors from life-table data. J Anim Ecol 44:85-114
- Sharpe PJH, Curry GL, DeMichele DW, Cole CL (1977) Distribution model of organism development times. J Theor Biol 66:21-38
- Siddiqui WH, Barlow CA (1973) Effects of some constant and alternating temperatures on population growth of the pea aphid, *Acyrthosiphon pisum* (Homoptera: Aphididae). Can Ent 105:145-156
- Southwood TRE (1978) Ecological Methods, 2nd Edition, Chapman & Hall, London
- Stinner RE, Gutierrez AP, Butler GD (1974) An algorithm for temperature-dependent growth rate simulation. Can Ent $106:519-524$
- Woodburn TL, Vogt WG, Kitsching RL (1978) Estimation of age of females in field populations of *Lucilia euprina* (Weidemann) (Diptera: Calliphoridae) using ambient temperature and solar radiation. Bull Ent Res 68:251-261

Received in January 1981