General Program for Determining Instar Distributions from Headcapsule Widths: Example Analysis of Mountain Pine Beetle (Coleoptera: Scolytide) Data

J. A. LOGAN, B. J. BENTZ, J. C. VANDYGRIFF, AND D. L. TURNER

Rocky Mountain Research Station, Logan Forestry Sciences Laboratory, USDA Forest Service,
860 North 1200 East, Logan, UT 84321

Environ. Entomol. 27(3): 555–563 (1998)

ABSTRACT Determining instar distributions from field-collected insect samples is a common problem in insect ecology. We describe a generalized computer program (HCAP) that determines an optimum separation rule based on the distribution of headcapsule widths. The program is initiated through visual interaction with the user to determine starting points for separation rules. HCAP then determines the optimum instar classification rules; estimates of means and standard deviations of headcapsule widths for each instar; estimated number in each instar; and probabilities of misclassification. Application of the program is illustrated by the analysis of 3 data sets of >10,000 observations of headcapsule widths obtained from field collections of larvae of the mountain pine beetle, *Dendroctonus ponderosae* Hopkins. This analysis resulted in finding significant differences between larvae collected from different hosts (ponderosa pine, *Pinus ponderosae* Lawson, and lodgepole pine, *Pinus contorta* Douglas variety latifolia Engelmann) and ecological circumstances (outbreak phase and geographic location). These results indicate that caution must be taken when extrapolating published results to new data or from one ecological situation to another.

KEY WORDS *Dendroctonus ponderosae,* headcapsule distributions, instar determination , Dyar's rule

DETERMINING INSTAR DESIGNATION for field collected insect samples is fundamental to both applied and basic research. Insect phenology and the resultant life-stage distribution is important in applied ecology because pest species are often vulnerable to control methods for only a restricted period in their life cycle. Control applications must be timed to coincide with this window of vulnerability to be effective. Achieving timing objectives typically requires a reliable method to determine the instar distribution of field collected samples. Determining instar distributions is required for life tables analysis, key factor analysis and other important ecological investigations. Frequently several instars are present at the same time and their size distributions overlap to some extent. The problem is, therefore, determining the appropriate instar for a sampled individual.

Growth of insect larvae is discontinuous, with the most measurable change in size occurring following molts (Chapman 1982). Heavily sclerotized structures, such as headcapsules, remain approximately the same size during a stadium, and can be used to differentiate larval instars (Daly 1985). Developing a frequency distribution of headcapsule width measurements in which individual peaks correspond to the

various instars was first used >100 yr ago (Dyar 1890), and has been successfully applied many times since. In this paper we describe an easy to use computer program (HCAP) for analysis of larval headcapsule data. This program is a generalization of techniques described by McClellan and Logan (1994) for analysis of gypsy moth, *Lymantria dispar* (L.), headcapsule data. Optimum instar classification rules, estimates of means and standard deviations of headcapsule widths for each instar, estimated number in each instar, and probabilities of misclassification are generated by HCAP. User interaction is required to visually determine initial guesses of separation points between instars (low points in the combined frequency distribution curve). Once this visual determination is made, the program automatically performs the analysis with no further user interaction required.

As an example application of this program, and also because it provides useful information in its own right, we apply HCAP for analysis of mountain pine beetle, *Dendroctonus ponderosae* Hopkins, headcapsule distribution data. In our investigations of mountain pine beetle population ecology, it has often been necessary to accurately determine instar distributions from field collected data. These studies have involved computation of stage dependent mortality (Cole 1975, 1981), climate and weather synchronization of adult emergence (Bentz et al. 1991, Logan et al. 1995, Schmid et al. 1993), and field validation of a life-systems model

(Bentz 1995). As a result, we have amassed large data sets that total >10,000 individual headcapsule measurements collected from a wide variety of habitats and climates. Classification rules for determining instar of field collected mountain pine beetle larvae, and the ecological implications of the results from this classification analysis, are discussed in the remainder of this article.

Materials and Methods

Description of HCAP. HCAP was developed using MATLAB (Math Works 1993) numerical computation and visualization software. MATLAB is an efficient matrix based language that provides a wide array of computational and visualization procedures. Program development is extremely fast and convenient using these powerful procedures. MATLAB programs are written as M-files, algorithmic pseudo-code that are much easier to understand than detailed C or FORTRAN code. It is therefore easy to further extend or adapt existing programs. M-files are written as ASCII code, which means they can be transported among different platforms. The major disadvantage is that an interested user must have a copy of MATLAB to run the M-file program, but that is mitigated by the fact that M-files can be compiled as either FORTRAN or C code, and distributed as executable files. Math Works also maintains its own Web site, which further augments dissemination of the program.

Data Input and Visualization. Data input to the program is from a simple flat file containing any number of headcapsule measurements. The only requirement is that datum be separated by 1 or more blanks. After the name of the file containing the data is entered, the user is provided the opportunity to iteratively determine an appropriate number of frequency classes used to represent the distribution of headcapsule widths. This is accomplished by visual inspection of the plotted histogram constructed by the user choosing the width of class intervals.

Estimation of Instar Distribution Parameters. Parameters required for optimum separation rules are estimated by assuming that the headcapsule distribution for each instar is normally distributed (Caltagirone et al. 1983, Got 1988, McClellan and Logan 1994, Schmidt 1996), and that the observed distribution is the sum of these individual normal distributions. The total frequency distribution is then

$$h_i = \sum_{i=1}^{n} a_i e^{-b_i (x - c_i)^2}. \qquad [1]$$

Following McClellan and Logan (1994), parameters are obtained by simultaneously fitting equation 1 to the combined data set, where h_i is the frequency of headcapsule width x. for the i-th of n instars, and parameters a_i, b_i, and c_i are fitted by nonlinear least squares. By assuming a normal distribution, estimates of mean headcapsule width ($\bar{x}_i = c_i$), variance in headcapsule width ($s_i^2 = 1/2b_i$), and estimated num-

ber of individuals in the sample that are in the i-th instar ($n_i = a_i \cdot (2\pi)^{1/2} / \Delta$, where Δ is the interval width of the frequency distribution), can be obtained. Simultaneously fitting these parameters accounts for potential overlap between instar distributions.

Classification rules can be obtained from areas of overlap between the individual instar frequency distributions. Consider the frequency distributions obtained from substituting the estimated parameters back into the assumed normal distributions, resulting in,

$$f_i = n \frac{\Delta}{\sigma \sqrt{2\pi}} e^{-(x - \mu)^2 / 2\sigma^2}. \qquad [2]$$

A classification rule follows by separating instars at the points of intersections of the individual instar frequency distributions. These intersections are obtained by numerically solving for $l_i = \text{root}(f_{i-1} - f_i)$ and $L_i = \text{root}(f_i - f_{i+1})$, where l_i is the lower limit for the i-th instar and L_i is the upper limit for the i-th instar.

The probability of misclassification of the i-th instar as the i-th $-$ 1 instar is given by,

$$P_l(l_i) = \frac{\displaystyle\int_{-\infty}^{l_i} f_i \, dx}{\displaystyle\int_{-\infty}^{+\infty} f_i \, dx}, \qquad [3]$$

and that of misclassifying the i-th instar as the i-th $+$ 1 instar is given by

$$P_L(L_i) = \frac{\displaystyle\int_{L_i}^{\infty} f_i \, dx}{\displaystyle\int_{-\infty}^{\infty} f_i \, dx}. \qquad [4]$$

The probability of misclassifying the i-th instar as some other instar is then $P_l(l_i) + P_L(L_i)$ and the probability of misclassifying some other instar as the i-th instar is $P_L(L_{i-1}) + P_l(l_{i+1})$. These probabilities are dependent on both the means and variances for the individual instar distributions and the number of individuals in each instar in the sample.

HCAP Program. HCAP is available as either a MATLAB M-file or as a compiled C-code executable version. Either version can be obtained by writing to J.A.L. There are different advantages to both versions. The M-file is meta-code that can be easily modified or expanded to meet specific needs. However, to do so, one needs a copy of MATLAB. The compiled C-version, on the other hand, requires no additional software link, but it is immutable.

MATLAB programs use 2 windows, the command window (Fig. 1), and a graphics window (Figs. 2 and 3). Both windows are interactive, and user response is communicated to the program by switching back-and-forth between them. The graphics and the command windows appear simultaneously on the screen.
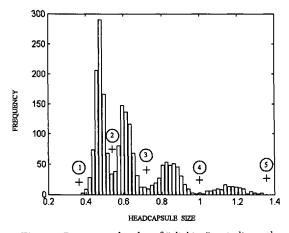
Fig. 1. Position and order of "clicking" to indicate the peaks in headcapsule distribution in data set II. Circled numbers were added to HCAP graphics window to indicate order of peak selection. + Indicates where the user would "click" to graphically input initial guesses of the optimum separation of instars.

Fig. 3. Zoom capabilities of HCAP. This figure is an enlargement of the region separating the 1st and 2nd instars. "Click-and-drag" with the left mouse button to zoom-in on a selected area. "Click" with the right mouse button to go back to the original graph. The lines on this figure are 1 the individual instar distributions; 2 the curves resulting from the original guesses; 3 the fitted normal series; 4 indicates the point of optimal (by McClellan/Logan criteria) separation of instars.

The primary function of the command window is to interact with the user to select the appropriate resolution (number of frequency classes) required to describe the distribution of headcapsules. This is accomplished by iteration until the user indicates satisfaction with the results. Selection of an appropriate frequency class width is important because of the smoothing effect that results. One extreme is that if frequency classes are too wide (too few classes selected), then existing, complex peaks may be smoothed over. Conversely, if frequency classes are too narrow (too many classes selected) resolution in peaks is lost, the extreme being classes are so narrow that each contains only one observation. In practice, we have found that 10 frequency classes per peak is a good starting point. A few choices greater and less than this value should quickly result in an acceptable resolution.
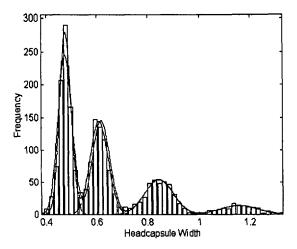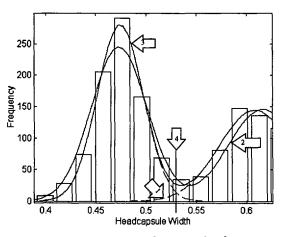
An appropriate selection of frequency classes is also important because the least-squares, normal series is fitted to the midpoints of the frequency classes. Choice of frequency interval width can, therefore, influence the final parameter selection and R^2 value.

The graphics window serves 3 purposes. First it provides for visual evaluation of resolution in frequency class representation of the headcapsule distribution. Second, it provides a way to input the original guesses of separation points (Fig. 1). Finally, the results of curve fitting are displayed in the graphics window (Fig. 2), along with the capabilities to "zoom" in on an interesting region of the distribution (Fig. 3).

HCAP Output. Example of HCAP output is listed in Table 1. Optimum separation rules are computed for both the 3 parameter function of equation 1, which includes the number of individuals in each instar and the standard normal distribution which is a function of the mean and standard deviation only. Descriptive statistics are computed from both the estimated normal distributions and from the classified, empirical data sets. Finally, the probabilities of miss-classification are computed from the summed 3 parameter functions.

Data Sources. Three sources of data were used for this analysis (Table 2). The headcapsule measurements that comprise data set I were collected between the years 1970 and 1991 from lodgepole pine, *Pinus contorta* Douglas variety *latifolia* Engelmann. Headcapsule measurements for these data resulted from numerous studies, and a subset (before 1983) of these measurements were published in Amman and Cole (1983). Larvae were collected from either building, epidemic or postepidemic populations from 6 national forests in western states. Collection times were typically once in the fall, once in the spring, and again in

Fig. 2. Fitted curves for the data set II. HCAP graphics are color coded, so interpretation is easier.

Table 1. Output produced by HCAP

>> HCAP *M-file name executed from MATLAB*

Enter data file name <= **93&95.dat**

How many frequency classes ("bars") for data plot? **50**

Are you happy with the plot? **y**

using the number of frequency classes you have selected (50), the
    frequency cell width is: 0.019400

Click on the places that separate instar distributions! *see Fig 1, this
    requires interactive input.*
Remember to click on BOTH the beginning and end of the
    distributions!
Press the [ENTER] key when you are done

The starting Sum-of-Squares = 73.766307
The elapsed time for curve fitting was 17.900000 seconds
The final Sum-of-Squares = 37.927493

The Crude R-square based on 50 frequency classes is = 0.920005
The Adjusted R-square based on 50 frequency classes is = 0.894743
Current plot held

Classification rules based on the 3 parameter normal distribution
3.506000e−001 < instar1 < 5.307405e−001
5.307405e−001 < instar2 < 7.072525e−001
7.072525e−001 < instar3 < 9.962242e−001
9.962242e−001 < instar4 < 1.359400e+000

Classification rules based on the 2 parameter normal distribution
3.506000e−001 < instar1 < 5.241438e−001
5.241438e−001 < instar2 < 6.951585e−001
6.951585e−001 < instar3 < 9.752071e−001
9.752071e−001 < instar4 < 1.359400e+000

Descriptive statistics estimated from the 3-parameter Normal
    Distribution

	Instar1	Instar2	Instar3	Instar4
Mean	0.474406	0.608250	0.842890	1.157059
Std. Dev	0.022754	0.035354	0.055988	0.073458
Number	824.994435	661.479508	390.662076	132.427271

Descriptive statistics estimated from empirical data classified by
    the 3 parameter Normal distribution

	Instar1	Instar2	Instar3	Instar4
Mean	0.472016	0.613165	0.839845	1.157926
Std. Dev	0.026429	0.034644	0.056358	0.076080
Number	863.000000	654.000000	388.000000	135.000000

Probabilities of missclassification based on the 3-parameter normal
    distribution

Instar	i as i−1	i as i+1	total
1	0.000000	0.006647	0.006647
2	0.014176	0.002553	0.016728
3	0.007704	0.003084	0.010788
4	0.014281	0.000000	0.014281

ans = *This is the "answer" matrix of fitted parameters produced by
    MATLAB*

2.8061e+002	9.6573e+002	4.7441e−001
1.4481e+002	4.0003e+002	6.0825e−001
5.4003e+001	1.5951e+002	8.4289e−001
1.3952e+001	9.2659e+001	1.1571e+000

>>

Bold font are user responses to HCAP queries, italicized font are
explanatory comments added to the output.

the middle of summer. Data set II also contains larval
headcapsule measurements collected from lodgepole
pine. However, these larvae were collected approxi-
mately once each month, with each consecutive sam-
ple taken from the same tree, thereby representing the
complete mountain pine beetle life-cycle for a specific
population. Larvae were collected from endemic pop-
ulations over a 2-yr period, in 3 different national

forests. Data set III includes measurements of larvae
collected over a 3-yr period from a building mountain
pine beetle population in ponderosae pine, *Pinus pon-
derosa* Lawson. These data, which came from only 1
national forest, were also collected approximately
once per month throughout the life-cycle.

All larvae were removed from phloem tissue in the
field, and stored in 95% ETOH until measurement in
the laboratory. Larval headcapsule width (to the near-
est 0.02 mm) was measured using a 50× power mi-
croscope and ocular grid-micrometer for data set I.
Measurements (recorded to the nearest 0.01 mm) for
data set II and data set III were made using a 25×
microscope with a Boeckeler digital eyepiece and
scaler.

Results

Means and Standard Deviations. Larval headcap-
sule widths for all 3 data sets ranged from 0.31 to 1.60
mm. The frequency distribution of widths indicate
four distinct peaks, representing 4 instars (Figs. 4–6).
Two previous studies also resulted in identification of
4 instars (Reid 1962, Amman and Cole 1983). Esti-
mated means, standard deviations, and number of in-
dividuals in each instar are listed in Tables 3–5.

There is a statistically significant difference in head-
capsule size due to instar, the different data sets, and
the interaction between them (PROC GLM, SAS In-
stitute 1989; Table 6). Using results from this analysis,
a least significant difference ranking indicates that all
means are significantly different from one another
with the exception of the 1st instar where the data set
I and data set III means were not statistically different
(Fig. 7). These differences could result from a variety
of causes (see *Discussion* section).

Our estimation of means and variances assumes that
the basic underlying distributions are normal. This
assumption is in keeping with previous instar separa-
tion techniques that take into account the problem of
overlapping distributions (Got 1988, Schmidt 1996). A
Kolmogorov–Smirnov test (Conover 1980) for nor-
mality resulted in inconsistent results. Data sets I and
III significantly deviated from the normal assumption
(Kolmogorov–Smirnov test statistic = 0.019, 0.043, re-
spectively; $P < 0.05$ with critical values computed from
a large sample approximation using $n = 7,270$ and
1,126, respectively), and normality was satisfied for
data set II (Kolmogorov–Smirnov test statistic = 0.02;
$P > 0.2$ with critical value computed from a large
sample approximation using $n = 2,040$). Lack of meet-
ing the normal assumption for data set I could result
from the broad temporal/spatial variability repre-
sented in this data. Regardless, the deviation from
normality is not a serious consideration for the analysis
of variance because of the innate robustness of the
technique and the high levels of significance obtained.
However, substantial deviation from normality would
effect the location of points separating instars.

Both means and standard deviations are an increas-
ing function of instar (Tables 3–5). The standard de-
viation is also an increasing function of the mean, a

Table 2. Larval data set descriptions

Data set name	Host	National forest(s), state	Yr	No. collections/yr
Data set I	Lodgepole pine	Cache, UT	1970–1975	3
		Wasatch, WY	1970–1975	
		Bridger, WY	1973–1974	
		Targhee, WY	1973–1975	
		Sawtooth, ID	1974	
		Teton, WY	1980–1991	
Data set II	Lodgepole pine	Cache, UT	1993–1994	10
		Targhee, WY	1993–1995	
		Sawtooth, ID	1993–1995	
Data set III	Ponderosa pine	Dixie, UT	1993–1994	10

result consistent with the general morphometric literature. This relationship can also be expressed by plotting the standard deviation as a function of the mean (Fig. 8). The linear relationship between standard deviation and the mean was highly significant ($F = 60.27$, df = 1, 10; $P < 0.01$).
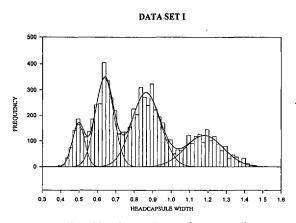
DATA SET I

Fig. 4. Plot of fitted normal series function to data set I. Histograms are frequency classes for observed headcapsule widths; light lines are individual instar distributions; heavy line is the fitted normal series.
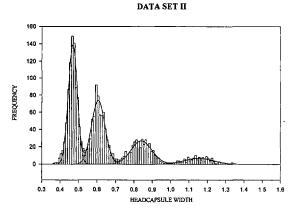
DATA SET II

Fig. 5. Plot of fitted normal series function to data set II. Histograms are frequency classes for observed headcapsule widths; light lines are individual instar distributions; heavy line is the fitted normal series.

Dyar (1890) hypothesized a geometric progression of headcapsule size with instar. This relationship results in $y = ab^x$, or following logarithmic transformation, $\ln y = \ln a + (\ln b)x$, where y is measured headcapsule size, x is instar, and a and b are estimated constants (Klingenberg and Zimmermann 1992). There has been a sustained interest in Dyar's rule over the past century because it has been interpreted as a simple way to determine the number of instars for an insect species from field collected data (e.g., representation of all instars results in a linear relationship
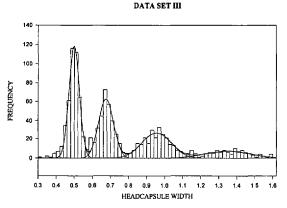
DATA SET III

Fig. 6. Plot of fitted normal series function to data set III. Histograms are frequency classes for observed headcapsule widths; light lines are individual instar distributions; heavy line is the fitted normal series.

Table 3. Estimated statistics for headcapsule distributions for data set I

Parameter	Instar 1	Instar 2	Instar 3	Instar 4
n (estimated)[a]	723	2152	2804	1577
n (classifeid)[b]	707	2169	2845	1549
Mean	0.493	0.638	0.866	1.189
Standard deviation	0.0355	0.0503	0.0785	0.105
Lower limit	—	0.543	0.730	1.03
Upper limit	0.543	0.730	1.03	—
$P(i$ as $i-1)$	—	0.0302	0.0415	0.0599
$P(i$ as $i+1)$	0.077	0.0332	0.0209	—
P(miss clasify)	0.077	0.0634	0.0624	0.0599

[a] The value for n is estimated from the parameter a of equation 1.
[b] The value for n is the sum of the classified data using the 3-parameter, normal series of equation 1.

Table 4.   Estimated statistics for headcapsule distributions for data set II

Parameter	Instar 1	Instar 2	Instar 3	Instar 4
n (estimated)[a]	825	661	391	132
n (classified)[b]	863	654	388	135
Mean	0.474	0.608	0.843	1.157
Standard deviation	0.0227	0.0354	0.056	0.0735
Lower limit	—	0.531	0.707	0.996
Upper limit	0.531	0.707	0.996	—
P(i as i − 1)	—	0.0142	0.0077	0.0143
P(i as i + 1)	0.0066	0.0025	0.0031	—
P(miss clasify)	0.0066	0.0167	0.0108	0.0143

[a] The value for n is estimated from the parameter a of equation 1.
[b] The value for n is the sum of the classified data using the 3-parameter, normal series of equation 1.

between the ln of headcapsule size and instar) (Hutchinson and Tongring 1984). Plots of the ln relationship and fitted linear regression lines are given in Fig. 9. All 3 fitted relationships are highly significant (F = 936.57, 579.86, 731.40 for data sets I, II, and III, respectively; df = 1, 2; $P < 0.01$). An analysis of covariance indicated that the adjusted means were significantly different (F = 15.10; df = 2, 8; $P < 0.01$), as were the regression coefficients (F = 8.57; df = 2, 6; $P < 0.05$). From visual examination, the slope of the fitted linear regressions (Dyar's constant) is similar for data sets I and II (0.29 and 0.30, respectively), whereas that for data set III (0.33) is greater.

Classification Rules. HCAP produces 2 classification rules, one derived from the 3-parameter function equation 1, and the other based on means and variances only. Both classifications are also listed in Tables

Table 5.   Estimated statistics for headcapsule distributions for data set III

Parameter	Instar 1	Instar 2	Instar 3	Instar 4
n (estimated)[a]	409	283	270	107
n (classified)[b]	446	309	267	104
Mean	0.510	0.677	0.955	1.37
Standard deviation	0.0270	0.0331	0.0979	0.113
Lower limit	—	0.588	0.760	1.175
Upper limit	0.588	0.760	1.175	—
P(i as i − 1)	—	0.0036	0.0227	0.0443
P(i as i + 1)	0.0019	0.0062	0.0124	—
P(miss clasify)	0.0019	0.0098	0.0351	0.0443

[a] The value for n is estimated from the parameter a of equation 1.
[b] The value for n is the sum of the classified data using the 3-parameter, normal series of equation 1.

Table 6.   Statistical analysis for headcapsule widths

Source	df	Sum square	Mean square	F	Pr $F >$
Data set	2	5.75	2.87	721.52	<0.0001
Instar	3	286.36	95.45	23,958	<0.0001
Data set * instar	6	2.87	0.49	120	<0.0001
Error	10,424	41.43	0.063		
Total	10,435	636.48			

Instars determined by classification rule derived from the 3 parameter, instar specific normal distribution. Sum squares are type III from SAS (SAS Institute 1989) procedure GLM.
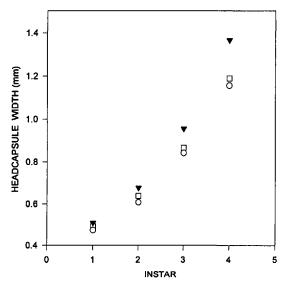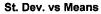

Fig. 7.   Estimated mean headcapsule widths for the 3 Mountain pine beetle data sets: □, data set I; ○, data set II; ▼, data set III.

3–5. As long as the survival schedule (l_x) is realistic (a monotonic decreasing function of age), and the variance is an increasing function of instar, the optimal separation between instars will always be shifted to the left of the apparent separation point (the low point in the headcapsule distribution). How far it is shifted will depend on the severity of mortality and the steepness of the increasing variance function.
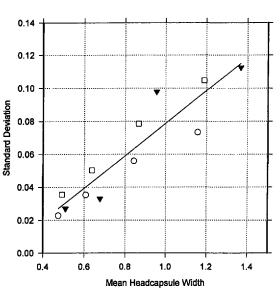

Fig. 8.   Estimated standard deviations of headcapsule widths as a function of mean: □, data set I; ○, data set II; ▼, data set III.
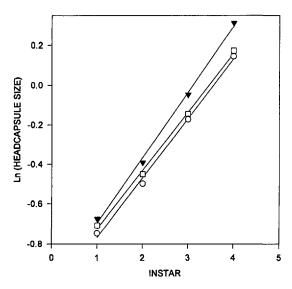
Fig. 9. Plots of fitted regression line to ln (headcapsule size) by instar. A linear relationship satisfies Dyar's rule, and indicates full representation of instars in the sample: □, data set I; ○, data set II; ▼, data set III.

Discussion

Means and Standard Deviations. The significant differences in mean headcapsule widths could result from a variety of sources. Adults reared from ponderosa pine are typically larger than those reared from lodgepole pine (Hay 1956, Wood 1963, Amman 1982), implying that larvae from ponderosa pine would also be larger. Size differences between hosts could result from both environmental and genetic factors. Ponderosa pine has thicker bark than lodgepole pine, and although they are in the same genus, chemical makeup (particularly defensive chemistry) among *Pinus* species is highly variable (Mirov 1961). These differences in both quantity and quality of the host food resource suggest potential significant nutritional differences that may result in size differences. Size differences could also simply be the result of geographic variation. Data set III was collected from trees at ≈37° 45' latitude, whereas both data sets I and II were collected farther north than 41° 56' latitude.

Population phase could also play a role in the observed size difference among larvae from the different data sets. Data set I is composite data collected over many years, but measurements were predominantly from outbreak phase populations. Conversely, data set II data were collected from endemic populations. At low population levels, mountain pine beetle typically attacks smaller, less vigorous trees, whereas larger populations are able to overcome the defenses of larger, healthier trees. Following this reasoning, data set I larvae may have come from more nutritious trees, resulting in larger larvae. Our results from the data set I analysis are similar to Amman and Cole (1983, p. 3). This result is not particularly surprising since their data is a subset of data set I. However, mean values

reported by Reid (1962) are also similar to those estimated from data set I, and are independent data. Reid also collected larvae from predominantly outbreak phase populations in lodgepole pine.

The observed differences in mean headcapsule widths indicate that caution must be taken when extrapolating published results to new data or from one ecological situation to another. An obvious question arises: Do the observed differences among data sets have a genetic basis, or are they simply due to local nutritional environment? We are conducting laboratory studies to more clearly elucidate the differences we observed.

Although the increasing variance with instar is an expected result, visual examination of Fig. 8 indicates that the variance of data set I is consistently greater than that for data set II with all the standard deviations for the data set I data lying above the fitted regression line and all those for data set II lying below. This difference is most likely a reflection of the larger geographic variability represented by data set I, and may also be an indication that greater variation in population phases that are represented in data set I. The pattern of variation in data set III exhibits no obviously consistent pattern.

Dyar's rule was satisfied for all 3 data sets. This is an indication that all instars are represented in the samples, a result that is consistent with the four observed peaks in the histogram plots (Figs. 5–7) and previously published data (Reid 1962, Amman and Cole 1983). In morphometric analysis, it is always advisable to have corroborative evidence regarding the number of instars for a particular species (see discussion below regarding smoothing functions).

Classification Rules. Classification rules used in HCAP are those described in equations 3 and 4. The rules minimize the pair-wise combined probability of misclassifying 2 consecutive instars. This is not the only classification rule that has been used. For example, Got (1988) chose to separate instars at the point that equalized the pair-wise probability of misclassifying 2 instars. Although HCAP does not compute Got's criteria, it can be easily calculated from information produced by HCAP and the maximum likelihood equations in Got (1988, p. 96).

HCAP produces 2 classification rules, one based on the 3 parameter equation 1, and the other on means and variances only. The probability of misclassifying individuals from any given sample is a function not only of the means and the variances, but also depends on the number of individuals in each instar. The 3 parameter function takes the number in each instar into consideration. For classification of a new individual larvae sampled from an unknown population, however, the situation is somewhat different. In this latter circumstance, if there is equal likelihood that the sampled individual is from any instar, then the optimum classification depends only on the appropriate means and variances. In this case, the 2 parameter rule should be used. Got (1988) also assumed equal (instar independent) sampling likelihood. The "best" decision rule, therefore, depends on the particular circum-

stances and applications intended for the classification rule. In cases where a large sample needs to be divided into representative instars (e.g., life history studies) it is probably best to always develop a new 3 (per instar) parameter classification scheme, particularly considering the multiple, and probably unknown, factors that can lead to headcapsule size differences between sampled populations. In situations where a small sample of individual needs to be identified to instar (e.g., pest management decisions in which the numbers in each instar are unknown and unestimatable), the 2 parameter rules given in HCAP or Got's criteria are recommended. An improved criteria could possible be developed by combining morphometric analysis with phenology models.

Lx Analysis. From equation 1, and its relationship to the normal curve, the interpretation of parameters b_i and c_i is straightforward as the inverse of twice the variance and the mean, respectively. However, there are 2 equally valid interpretations of the parameter a_i. First, from examination of equation 1, it is apparent that a_i is a parameter that scales the frequency distribution of the i-th instar to its maximum. Second, by relating a_i to the normal curve, n_i is an independent estimate of the number of individuals in the total sample that belong to the i-th instar corrected for overlapping distributions with other instars, and is estimated as

$$n_i = a \frac{\sigma \sqrt{2\pi}}{\Delta}. \qquad [5]$$

If certain conditions are met, the relationship can be used to develop survival schedules and a decision rule that is independent of sample size.

HCAP. In our experience, HCAP has provided an easy to use, efficient way to determine instars separation rules from large data samples. Convergence properties of equation 1 and the least squares criteria appear to be robust. HCAP, however, does not determine the number of peaks or instars that compose a complex, composite distribution. That determination is left to the user who determines the number of terms to include in the normal series by selecting the initial separation points. HCAP then fits the normal series of equation 1 that includes the identified number of terms in the series. Other commercial programs have been applied to the problem of separating headcapsule distributions by assuming a composite normal distribution. For example, Schmidt (1996) applied PeakFit (Jandel 1990) for morphometric analysis of spruce budworm headcapsule distributions. PeakFit is similar to HCAP in that the normal series of equation 1 is fitted by nonlinear least squares, however, a statistical analysis of results is performed to help determine the number of terms in the series that are required to adequately describe a composite distribution. The situation with spruce budworm is more confusing than our application with mountain pine beetle, supernumerary instars and effects of parasitism adding complexity to the budworm analysis. Schmidt reported good results with PeakFit in resolving this complex

situation. Although the distributions for mountain pine beetle appear to be straightforward, we have found that HCAP is capable of producing satisfactory results for complex morphometric distributions as long as visual examination can provide good initial estimates of separation of peaks. An approach that might work well for complex composite curves would be application of a program like PeakFit to determine the number of instars represented, and then apply HCAP to compute the optimum separation rules and probabilities of misclassification.

PeakFit also applies a smoothing algorithm before fitting the normal series. HCAP provides a smoothing function by iteratively allowing the user to determine the number of frequency classes used to describe the data. As an alternative, we also applied a 5-term moving average smoothing algorithm combined with first differences of the moving average to determine points of inflection between instars. The moving average approach worked well for our 3 data sets, but did not result in a noticeable improvement over visual inspection. In our experience, visual separation works well, and it has required only a few iterations to obtain satisfactory results. Visual qualitative determination also effectively addresses the possibility of inadvertently ignoring a peak that could be smoothed over by an automatic algorithm.

Acknowledgments

References Cited

Amman, G. D. 1982. Characteristics of mountain pine beetles reared on four pine hosts. Environ. Entomol. 11: 590–593.

Amman, G. D., and W. E. Cole. 1983. Mountain pine beetle dynamics in lodgepole pine forests. Part II: population dynamics. U. S. Dep. Agric. For. Serv. Gen. Tech. Rep. INT-145.

Bentz, B. J. 1995. Ecological adaptations of the mountain pine beetle to cold temperatures, pp. 67–85. *In* F. P. Hain, S. M. Salom, W. F. Ravlin, T. L. Payne, and K. F. Raffa [eds.], Proceedings of the Joint IUFRO Conference: Behavior, Population Dynamics and Control of Forest Insects, 6–11 February 1994, Maui, HI. The Ohio State University, Wooster, OH.

Bentz, B. J., J. A. Logan, and G. D. Amman. 1991. Temperature dependent development of the mountain pine beetle (Coleoptera: Scolytidae), and simulation of its phenology. Can. Entomol. 123: 1083–1094.

Caltagirone, L. E., W. Getz, and D. W. Meals. 1983. Head capsule width as an index of age in larvae of navel orangeworm, *Amyelois transitella* (Lepidoptera: Pyralidae). Environ. Entomol. 12: 219–221.

Chapman, R. F. 1982. The insects: structure and function. Harvard University Press, Cambridge, MA.

Cole, W. E. 1975. Interpreting some mortality factor interactions within mountain pine beetle broods. Environ. Entomol. 4: 97–102.

1981. Some risks and causes of mortality in mountain pine beetle populations: a long-term analysis. Res. Popul. Ecol. 23: 116–144.

Conover W. J. 1980. Practical nonparametric statistics. Wiley, New York.

Daly, H. G. 1985. Insect morphometrics. Annu. Rev. Entomol. 30: 415–438.

Dyar, H. G. 1890. The number of molts of lepidopterous larvae. Psyche (Cambr.) 5: 420–422.

Got, B. 1988. Determination of instar of the European corn borer (Lepidoptera: Pyrtalidae) based on a distribution model of headcapsule widths. Ann. Entomol. Soc. Am. 81: 91–98.

Hay, C. J. 1956. Experimental crossing of mountain pine beetle with Black Hills beetle. Ann. Entomol. Soc. Am. 49: 567–571.

Hutchinson, A. M., and N. Tongring. 1984. The possible adaptive significance of the Brooks—Dyar-rule. J. Theor. Biol. 106: 437–439.

Jandel. 1990. PeakFit peak analysis software; reference manual, version 2. Jandel, Corte Madera, CA.

1994. SigmaPlot scientific graphing software: user's manual. Jandel, San Rafael, CA.

Klingenberg, C. P., and M. Zimmerman. 1992. Dyar's rule and multivariate allometric growth in nine species of waterstriders (Heteroptera: Gerridae). J. Zool. (Lond.) 227: 453–464.

Logan, J. A., P. V. Bolstad, B. J. Bentz, and D. L. Perkins. 1995. Assessing the effects of changing climate on mountain pine beetle dynamics. pp. 86–108. In F. P. Hain, S. M. Salom, W. F. Ravlin, T. L. Payne, and K. F. Raffa [eds.], Proceedings of the Joint IUFRO Conference: Behavior,

Population Dynamics and Control of Forest Insects, 6–11 February 1994, Maui, HI. The Ohio State University, Wooster, OH.

Math Works. 1993. MATLAB reference guide. Math Works, Natick, MA.

McClellan, Q. C., and J. A. Logan. 1994. Instar determination for the gypsy moth (Lepidoptera: Lymantriidae) based on the frequency distribution of headcapsule widths. Environ. Entomol. 23: 248–253

Mirov, N. 1961. Composition of gum terpentines of pines. U.S. Dep. Agric. For. Serv. Tech. Bull. 1239.

Reid, R. W. 1962. Biology of the mountain pine beetle, Dendroctonus monticolae Hopkins, in the eastern Kootenay region of British Columbia. I. Life cycle, brood development, and flight periods. Can. Entomol. 94: 532–538.

SAS Institute. 1989. SAS/STAT users guide, version 6, 4th ed., vol. 2. SAS Institute, Cary, NC.

Schmid, J. M., S. A. Mata, W. K. Olsen, and D. D. Vigil. 1993. Phloem temperatures in mountain pine beetle-infested ponderosa pine. U.S. Dep. Agric. For. Serv. Res. Note RM-521.

Schmidt, F. H. 1996. Instar discrimination of field-collected larvae through analysis of frequency distribution curves of head capsule widths using the program PeakFit. U.S. Dep. Agric. For. Serv. Res. Pap. PNW-RP-492.

Wood, S. L. 1963. A revision of the bark beetle genus Dendroctonus Erichson (Coleoptera: Scolytidae). Great Basin Nat. 23: 1–117.