

Abstract—We consider estimation of mortality rates and growth parameters from length-frequency data of a fish stock and derive the underlying length distribution of the population and the catch when there is individual variability in the von Bertalanffy growth parameter L_∞ . The model is flexible enough to accommodate 1) any recruitment pattern as a function of both time and length, 2) length-specific selectivity, and 3) varying fishing effort over time. The maximum likelihood method gives consistent estimates, provided the underlying distribution for individual variation in growth is correctly specified. Simulation results indicate that our method is reasonably robust to violations in the assumptions. The method is applied to tiger prawn data (*Penaeus semisulcatus*) to obtain estimates of natural and fishing mortality.

Maximum likelihood estimation of mortality and growth with individual variability from multiple length-frequency data

You-Gan Wang

CSIRO Mathematical and Information Sciences
65 Brockway Road
Floreat Park
Western Australia 6014, Australia
E-mail address: You-Gan.Wang@csiro.au

Nick Ellis

CSIRO Marine Research
P.O. Box 120
Cleveland, Queensland 4163, Australia

Estimation of growth and mortality is fundamental in fisheries because stock assessment and management rely on these population parameters. Length-frequency-based methods become important when aging techniques are either not possible or very expensive. Existing methods such as that of Beverton and Holt (1956) assume that recruitment is continuous and constant throughout the year, which leads to a population with an exponentially distributed age structure. Existing modifications to Beverton and Holt's method comprise some simple recruitment patterns or distributions (Ssentongo and Larkin 1973; Ebert 1980; Hoenig 1987; Wetherall et al. 1987). As pointed out by Vetter (1988), the existing methods for estimating mortality in the literature have strong limitations and disadvantages. In particular, they require the following assumptions:

- 1) each individual follows the same von Bertalanffy growth curve;
- 2) the recruitment is either continuous and constant throughout the year (as in Beverton and Holt [1956] and Wetherall et al. [1987]) or is a pulse function (as in Hoenig [1987]);
- 3) the total instantaneous mortality rate, z , is constant.

As pointed out by Sainsbury (1980), it is more realistic to allow individual variability in growth. For example, using tag-recapture data, Wang et al. (1995) found substantial individual variability for the tiger prawn species *P. semisulcatus*.

Estimation of mortality relies on the distribution of the lengths, which is determined by the age distribution, mortality rates, and the individual variability in growth rates. If individual variability in growth is ignored, an inappropriate length distribution will be generated, leading to biases in parameter estimates. It is also biologically interesting to quantify the individual variability in growth, which has important implications in fisheries management. Although it is well understood that variability leads to increased uncertainty in estimates, it is less well recognized (among the fisheries community) that variability can also lead to bias. Wang and Ellis (1998) analyzed the effect of ignoring individual variability in a simplified context of constant recruitment and a single length-frequency record. They found that, in the presence of individual variability, existing methods gave positively biased parameter estimates. More details about the background can be found in Ebert (1973), Askland (1994), and Wang and Ellis

Manuscript submitted 5 March 2004
to the Scientific Editor's Office.

Manuscript approved for publication
9 November 2004 by the Scientific Editor.
Fish. Bull. 103:380–391 (2005).

(1998). See DeLong et al. (2001) for alternative approaches to length-frequency data where individual variability is taken into account.

In our study, we develop a new framework for analyzing length-frequency data. In particular, we incorporate 1) individual variability in growth parameters; and 2) an arbitrary recruitment function. The model is flexible enough to incorporate various sizes at recruitment and a fishing selectivity function. However, we did not use these aspects in the analysis of tiger prawn data. Some analytical expressions are derived for these generalizations. A maximum likelihood approach is developed for estimation of mortality and growth parameters. Separation of fishing mortality from natural mortality is possible only when there is substantial contrast in the effort pattern. We also require a known recruitment pattern, and sampling times are spread out so that the length-frequency data will contain information on growth and mortality. Simulation studies are carried out to determine the performance of the method. The simulated data are generated from the recruitment pattern of the brown tiger prawn (*Penaeus esculentus*) in the northern prawn fishery of Australia. Finally we apply the maximum likelihood method to length-frequency data from grooved tiger prawn data (*P. semisulcatus*) in the northern prawn fishery of Australia.

Materials and methods

The model

We assume that the growth of individuals follow a von Bertalanffy curve so that the length at age a (relative to some origin t_0) is given by

$$L(a) = L_\infty(1 - e^{-ka}). \tag{1}$$

In this study, age is always defined to be relative to t_0 , i.e. t_0 is absorbed into a for the purpose of identifiability. We will consider estimation of (k, l_∞) only because t_0 is not estimable from length-frequency data with aging data. Note that this does not mean t_0 is assumed to be 0. To provide a general treatment we relax each of the assumptions mentioned in the introduction. First we relax assumption 1 by letting the maximum length, L_∞ , vary within the population. We denote the density function of L_∞ as $p(x)$, which has a mean of l_∞ and a variance of σ^2 . It is possible that recruits to the fishery have a range of sizes. To allow for this range we let the size at recruitment, L_0 , be a random variable with density function $u(s)$. In practice, one may be able to use information from other studies (such as subadult abundance) to arrive at an approximate parametric form for $u(s)$.

If $f_t(l)$ is the probability density function of L at time t , then

$$f_t(l) = \int_0^\infty \int_l^\infty p(x|L_0=s) f_t(l|L_\infty=x, L_0=s) u(s) dx ds, \tag{2}$$

where $f_t(l|L_\infty=x, L_0=s)$ is the conditional probability density function of L at time t when L_∞ is known to be x and the size at recruitment is s . Note the lower limit of the inner integral is l because L_∞ cannot be less than an individual's length.

Let the age (again, relative to t_0) at recruitment of an individual be A_0 . From Equation 1, we have age a at length l is $a = -k^{-1} \log(1 - l/L_\infty)$ and hence the conditional distribution, $f_t(l|L_\infty=x, L_0=s)$, which may be written as $f_t(l|x, s)$ for brevity, can be expressed by using the conditional distribution of age $ht(a|L_\infty=x, A_0=a_0)$ (see Wang et al., 1995), as

$$f_t(l|x, s) = \frac{1}{k(x-1)} h_t(-k^{-1} \log(1 - l/x) | x, a_0). \tag{3}$$

We now generalize assumptions 2 and 3 by introducing the intensity function of recruitment, $r(t)$, and the total instantaneous mortality, $z(t)$, which are arbitrary functions of time t . The total mortality would depend on time through the fishing mortality component F , where $z(t) = M + F(t)$ and M is the constant natural mortality.

The age distribution satisfies

$$h_t(a|L_\infty=x, A_0-a_0) \propto \exp\left[-\int_{a_0}^a z(t-a+y) dy\right] r(t-a+a_0). \tag{4}$$

This equation states that the density of individuals of age a is proportional to the intensity of recruitment at the time when these individuals were recruited, namely $t-a+a_0$, multiplied by a reduction factor due to mortality over the intervening period. We therefore have

$$h_t(a|x, s) = h_t\left(a|L_\infty=x, A_0=-k^{-1} \log(1-s/x)\right) \propto \exp\left(-\int_{-k^{-1} \log(1-s/x)}^a z(t-a+y) dy\right) r\left(t-a-k^{-1} \log(1-s/x)\right) \tag{5}$$

and Equation 3 becomes (after substituting for a and shifting the dummy variable y)

$$f_t(l|x, s) \propto \frac{1}{x-1} \exp\left(-\int_{t-k^{-1} \log\left(\frac{x-s}{x-1}\right)}^t z(y) dy\right) r\left(t-k^{-1} \log\left(\frac{x-s}{x-1}\right)\right). \tag{6}$$

Let us consider the case of fixed recruitment length, i.e., $L_0=l_0$, and define a parameter vector, β , consisting of (k, l_∞, s) , and other parameters quantifying mortality and catchability. Equation 2 then reduces to a single integral over x ,

$$f_t(l|\beta) \propto \int_l^\infty p(x) \exp\left(-\int_{t-k^{-1} \log\left(\frac{x-l_0}{x-1}\right)}^t z(y) dy\right) r\left(t-k^{-1} \log\left(\frac{x-l_0}{x-1}\right)\right) \frac{dx}{x-1}. \tag{7}$$

A more convenient form for computation arises after changing the integration variable from the asymptotic length x to time since recruitment, $t-a+\alpha_0$,

$$\tau = k^{-k} \log\left(\frac{x-l_0}{x-1}\right). \tag{8}$$

The expression (Eq. 7) then becomes

$$f_t(l|\beta) \propto \int_l^\infty p(x(\tau)) \exp\left(-\int_{t-\tau}^t z(y) dy\right) r(t-\tau) \frac{d\tau}{1-e^{-k\tau}}. \tag{9}$$

In the special case of constant recruitment, i.e., $r(t)=1$, and constant mortality, $z(t)=z$, $f_t(l|\beta)$ becomes independent of time as first obtained by Powell (1979).

Maximum likelihood estimation

Let $p_{ij}(\beta)$ be the expected proportion of individuals in the i^{th} length class (l_{i-1}, l) on the j^{th} occasion, where $i=1, 2, \dots, N$; and let n_{ij} be the corresponding observed numbers. The value of $p_{ij}(\beta)$ can be obtained from the density function $f_t(l;\beta)$ given by Equation 2. Thus

$$p_{ij}(\beta) = \frac{\int_{l_{i-1}}^{l_i} f_j(l;\beta) dl}{\int_{l_0}^{l_N} f_j(l;\beta) dl}, \tag{10}$$

in which $f_j(l;\beta)$ is the (unnormalized) density function on the j^{th} occasion. Under a multinomial model, estimation of the parameter vector β relies on the procedure

$$\text{maximize } \sum_{i,j} n_{ij} \log p_{ij}(\beta) \text{ with respect to } \beta. \tag{11}$$

The sum is the log-likelihood function up to a constant independent of the parameters. The probability, p_{ij} , can be approximated as $f_j(l_{i+1/2})/\sum_i f_j(l_{i+1/2})$, which is the normalized value of the density function for the j^{th} occasion at the midpoint of the i^{th} length class.

If sampling effort is known and expected catch is assumed to be a known function of effort and population abundance, the log-likelihood function in Equation 11 can be easily modified to incorporate effort information. For example, if the total number of individuals on each occasion, $n_j = \sum_i n_{ij}$, is assumed to follow a Poisson model with overdispersion parameter v , the log-likelihood function becomes

$$\sum_{i,j} n_{ij} \log p_{ij}(\beta) + v \sum_j \left\{ n_j \log \lambda_j(\beta) - \lambda_j(\beta) \right\}, \tag{12}$$

where $\lambda_j(\beta)$ is the expected total number in the sample on the j -th occasion and depends on effort. One way to model this dependence is $\lambda_j(\beta) = \phi p_j(\beta) e_j$, where e_j is the

sampling effort, ϕ is the total abundance index over all occasions; and p_j is the expected proportion of individuals on the j^{th} occasion (i.e., the relative abundance), so that ϕp_j is the expected catch per unit of effort. In this case we can obtain the maximum likelihood estimate of ϕ as $\sum_j n_j / \sum_j e_j p_j$. The probability, p_j , can be approximated as $\sum_i f_j(l_{i+1/2}) / \sum_{i,j} f_j(l_{i+1/2})$. Here v is introduced to allow for overdispersion in the Poisson model. It plays a weighting role for the two terms in Equation 12, and the second summation can be regarded as auxiliary information. If n_j is assumed to follow a Poisson distribution exactly, we have $v=1$.

In our simulation and tiger prawn studies we specify a case of fixed, known recruitment length, l_0 , and $f_j(l;\beta)$ is obtained from Equation 7 or 9. For definiteness we set the constant of proportionality implicit in these equations to one.

The integrals in Equations 7 and 9 present some subtleties for their evaluation, so that some details of the numerical implementation might be of interest. For the simulation study we used Equation 7. The integral was performed on an l -dependent grid of 41 and 81 quantiles of the L_∞ distribution $p(x)$ and then improved upon by using the Richardson extrapolation. Note that there is an apparent singularity at $x=1$. However, by decomposing the mortality into a mean and deviation term, $z(y) = \bar{z} + z(y) - \bar{z}$, we find that the factor involving mortality is proportional to $(x-l)^{\bar{z}/k}$. Hence the integrand is proportional to $(x-l)^{\bar{z}/k}$, and, because $\bar{z}/k - 1 > -1$, the singularity is integrable (i.e., the integral is finite). We used a quadrature scheme designed for integrands of the form $(x-l)^\xi f(x)$, $\xi > -1$, to perform the integral in the neighborhood of $x=1$.

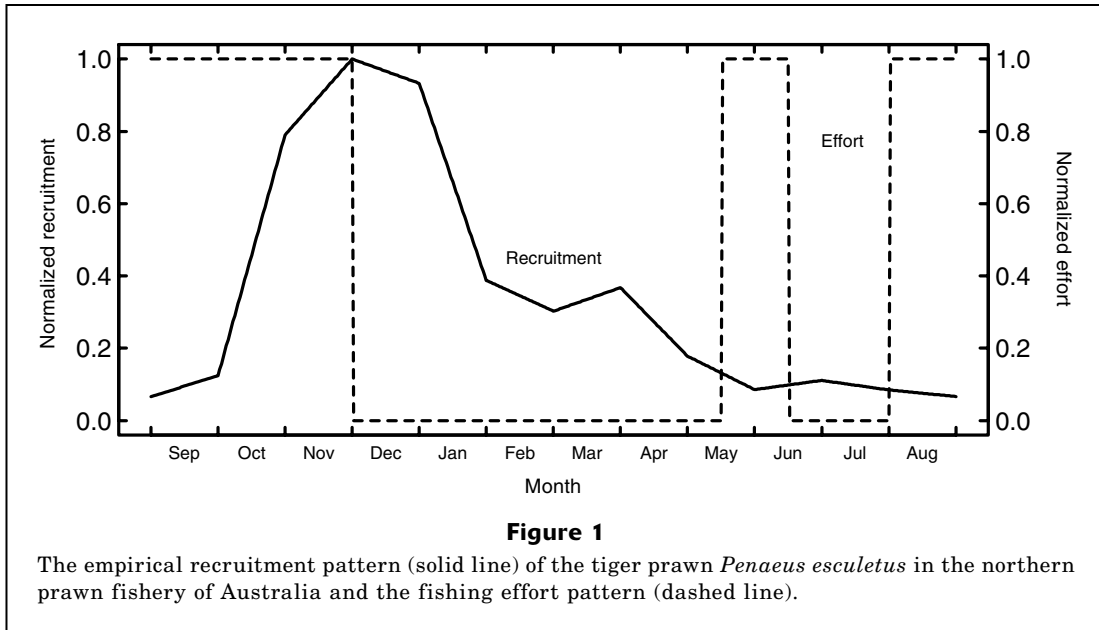
For the tiger-prawn study we used Equation 9. The integral was performed on uniform grids of 41 and 81 points over the interval $\tau \in (0, 1.5)$ years and, as before, was improved by using the Richardson extrapolation. We used our knowledge that tiger prawns live for about 18 months to determine the upper limit of integration. Note that despite appearances, this integral contains no singularity because $x(\tau) \rightarrow \infty$ as $\tau \rightarrow 0$, and therefore the factor $p(x(\tau)) / (1 - e^{-k\tau}) \rightarrow 0$. The effort integral within the integrand was computed by linear interpolation between cumulative totals of the weekly effort.

The prototype implementation of our maximum likelihood method was written in S-plus software (Lucent Technologies) by using the optimizer "nlminb." However, to improve the performance for a large number of simulations, the program was recoded in C by using Powell's optimization routine with numerical derivatives (Press et al., 1992). The C code and some relevant reports are available on request.

Results

Simulation studies

We simulated length-frequency data based on the recruitment pattern of tiger prawns *P. esculentus* in



the northern prawn fishery of Australia. This pattern has been derived from experimental trawls in which the number of individuals in the lowest length class are counted (Wang and Die, 1996). We assume the recruitment and effort patterns are the same in each year (Fig. 1). The effort pattern (dashed line) consists of two constant-fishing periods: 15 May to 15 June, and 1 August to 1 December. The unit of effort, E , depends on the unit of catchability, q , because the fishing mortality $F=qE$ must have unit yr^{-1} : therefore we let $E=1$ during the fishing season. Note that the proportion of the year that is fished is $\int E(t)dt=5/12$.

The growth component of our models has $l_{\infty}=40$ mm and $k=3yr^{-1}$; the instantaneous natural mortality is $M=2yr^{-1}$; and the instantaneous fishing mortality, F , during the fishing season is $4yr^{-1}$ (i.e., $q=4$, because in our units, $F=q$). The resulting annual mortality, $Z=\int z(t)dt=M+q\int E(t)dt=2+4\times 5/12=11/3$. The values for mortality come from Somers and Wang (1996). We assume that all recruits have length 19.5 mm. The L_{∞} distribution is normal (standard deviation 4 mm) but is truncated at 19.5 mm. The truncated normal distribution at l_0 is simply a conditional normal distribution conditional on being greater than l_0 .

We generate twelve length-frequency data sets, one for the beginning of each month. We choose a monthly time interval because the data from our case study in the next section were sampled at roughly monthly intervals. In addition, because the recruitment pattern is periodic it is sufficient to analyze one year of data.

We obtain each monthly length-frequency data set by taking a sample of size 1000 from the theoretical length distribution $f_i(l)$ given by Equation 6, which depends on the recruitment pattern, the effort pattern, and the distribution of L_{∞} . That is, for each of the 12 time points t , we evaluate numerically the right-hand

side of Equation 6 over a set of finely spaced l values (i.e., every 0.25 mm), aggregate the $f_i(l)$ to 1-mm intervals and finally normalize the function by dividing by the sum of $f_i(l)$. This results in an array of probabilities for an individual's length in each 1-mm interval. It is then straightforward to sample from the corresponding multinomial distribution.

We then obtain parameter estimates from the twelve months of simulated data. The process is repeated 100 times to provide a reasonable estimate of the sampling variance of the parameters. In practice, (k, l_{∞}) can often be estimated from a different study. We therefore consider two models. In model 1, we assume all five parameters are unknown, and, in model 2, we assume that l_{∞} and k are known and we estimate M , F , and σ . It is also common practice (e.g., Sullivan, 1992) to assume that M is known and to estimate the remaining parameters; this is the case in our model 3.

The results are summarized in Table 1. All the parameters are quite well estimated, even for model 1. Estimates of both natural mortality and fishing mortality are quite reliable when growth parameters are assumed known. There is also a modest reduction in the standard deviation when (k, l_{∞}) are assumed known.

We have also tested for robustness by performing the estimation process on data generated from a log-normal distribution. The results are shown in Table 1. For model 1 the estimates of M and F have a larger and opposite bias, whereas the absolute bias for Z is somewhat smaller. Model 2 improves the estimates dramatically, despite the fact that an incorrect distribution (the truncated normal) is being used in the model. Note that the variation in the estimates of total annual mortality, Z , is somewhat less than that for F and M ; this is because F and M are highly negatively correlated (typically 94%). In model 3 the estimate of

Table 1

Mean parameter estimates and standard deviations (in parentheses) for simulated tiger prawn (*Penaeus esculentus*) data. The model assumes an underlying truncated normal L_∞ distribution. The data are generated from two underlying L_∞ distributions: the truncated normal and the lognormal. With model 1 all parameters are unassumed to be unknown; with model 2 (k, l_∞) are assumed to be known; with model 3 M is assumed to be known.

Model	k	l_∞	σ	Z	M	F
Underlying truncated normal distribution						
True	3	40	4	3.67	2	4
1	2.99 (0.05)	40.00 (0.19)	4.02 (0.08)	3.65 (0.05)	1.98 (0.15)	3.99 (0.34)
2	3	40	4.01 (0.07)	3.65 (0.04)	2.00 (.11)	3.95 (0.28)
3	2.99 (0.05)	40.02 (0.15)	4.01 (0.07)	3.65 (0.05)	2	3.95 (0.12)
Underlying lognormal distribution						
True	3	40	4	3.67	2	4
1	3.02 (0.07)	39.53 (0.22)	4.28 (0.08)	3.53 (0.05)	1.51 (0.16)	4.84 (0.35)
2	3	40	4.14 (0.07)	3.62 (0.04)	1.93 (0.11)	4.05 (0.28)
3	2.96 (0.06)	39.92 (0.17)	4.16 (0.07)	3.57 (0.05)	2	3.76 (0.11)

F is negatively biased, but once again the standard deviation is reduced.

Application to tiger prawns (*P. semisulcatus*)

The data for this application consist of a six-year sequence of experimental length-frequency data from the trawling region around Albatross Bay in the eastern Gulf of Carpentaria, Australia. The data consist of catches of tiger prawns from 11 mm to 59 mm (carapace length) for each of 69 times ranging from March 1986 to March 1992. The catches from several stations covering the trawling region at each time (over a few consecutive days) are aggregated. Sampling was done roughly every lunar month.

We use the catch data for the smaller size classes to obtain two types of recruitment patterns: the *aperiodic* pattern and the *quasiperiodic* pattern. The aperiodic pattern is constructed by summing over all individuals with length 21 mm or less for each occasion. The resulting sequence of plotted time points is then joined up by straight lines. The quasiperiodic pattern is generated from the aperiodic pattern by averaging corresponding points across years to give a single annual pattern. The pattern for all six years is generated from the annual pattern by applying, for each biological year, a scale factor that is found by averaging the catch over all size classes within the year. The start of the biological year is defined as the time when the annual pattern reaches its minimum (see Fig. 2).

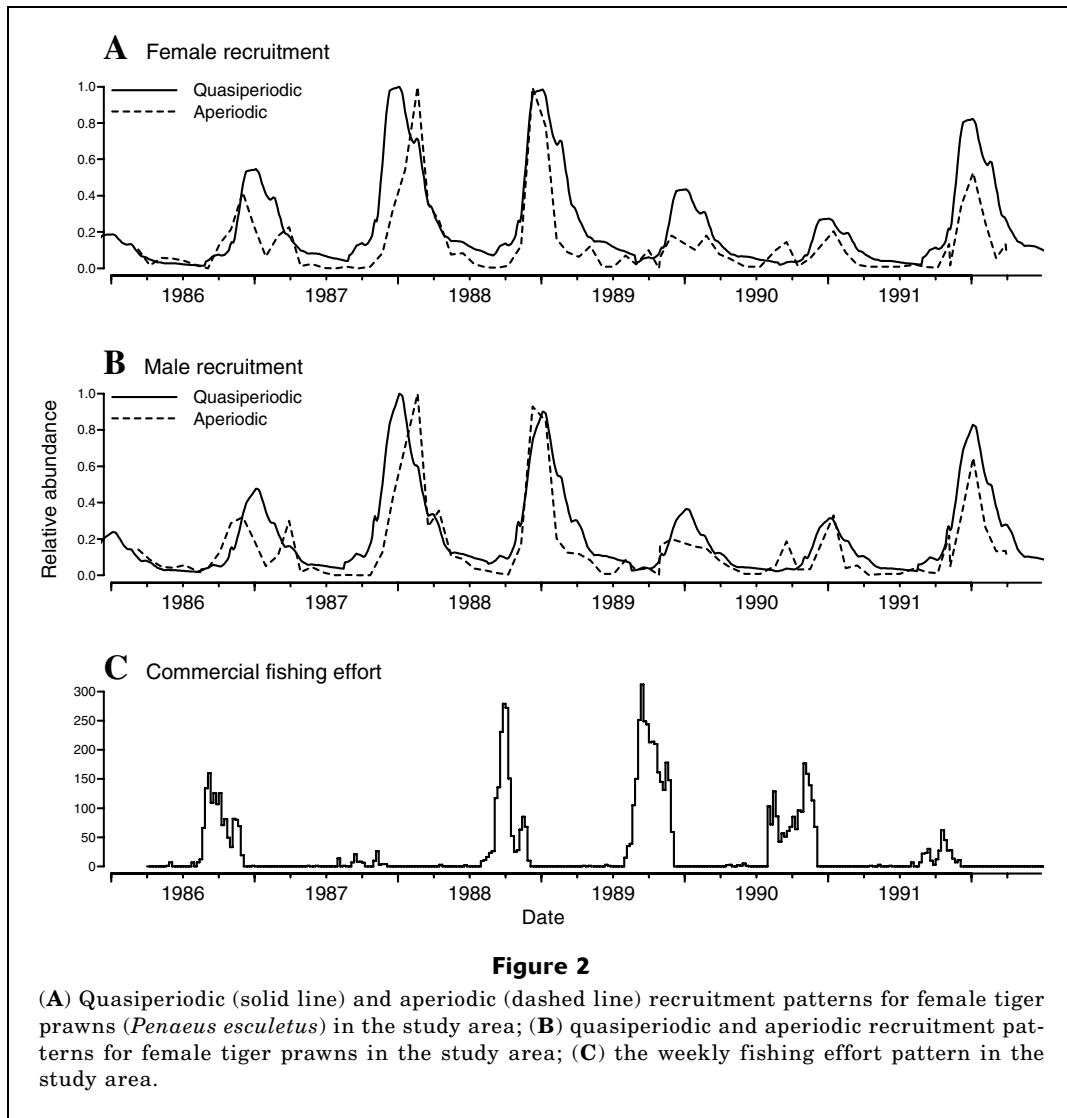
The effort pattern comes from commercial log books collected from fishermen for the period from 1986 to 1992 in the area. Effort is measured in boat-days (see Fig. 2). There is substantial contrast in the effort both within years (due to seasonal closures) and across years. This contrast may allow us to separate fishing mortality from natural mortality.

The instantaneous fishing mortality $F(t)$ is assumed to be $qE(t)$. The mean total mortality $Z = M + q\bar{E}$, where \bar{E} is the mean effort over the study period. Given the results of the simulation study, we expect the parameter Z may be more reliably estimated than either M or q , whose estimates are negatively correlated.

We further assume that the L_∞ distribution is a truncated normal distribution. This choice is based on the shape of the observed length distribution from July to September, the period when this distribution should approximate the asymptotic length distribution. The truncated normal distributions are then reparameterized in terms of the mean l_{∞^*} and variance σ_*^2 of this underlying normal distribution. It is more convenient to use these parameters than the mean l_∞ and variance σ^2 of the truncated normal distribution. Note that l_∞ is always larger than l_{∞^*} and σ is always less than σ_* . However, in this application the two sets of parameters are nearly interchangeable because over the range of estimated values l_∞ exceeds l_{∞^*} by at most 0.5 mm and σ_* exceeds σ by at most 0.6 mm (see Table 2).

We define a recruit to be an individual with length l_0 , which can be chosen at discretion. We examine a range of candidate values of l_0 between 19.5 mm and 27.5 mm, to find out which values provide the most suitable definition of recruitment for this data set, i.e., that which leads to the least violation of model assumptions.

In our application the recruitment pattern was derived from size classes 21 mm or less. If we use this pattern at say 23.5 mm then we need to shift the pattern slightly to later times. It is not apparent to what degree we should shift the pattern; therefore we shall estimate the degree of shift. We call this parameter the *lag*. We expect the lag to increase with l_0 . Also note that the derived recruitment pattern is an average over different size classes and hence it is an average over different times. The absolute timing of the pattern is



therefore uncertain and so the lag parameter adopts the role of estimating this uncertainty.

We do have sampling effort information, so that it would be reasonable to consider incorporating into the likelihood the Poisson term for the total catch as mentioned in section 3. Information on total catch per occasion would improve estimates of mortality. However, preliminary analysis found that there was a mismatch of the expected total catch with the observed total catch. Therefore, it appears to be unrealistic to assume that the catch is proportional to the sampling effort. In the subsequent data analysis we use the form of the log-likelihood in Equation 11, which uses the shape of the observed distribution and takes the total catch as given.

We have estimated all the parameters k , l_{∞} , σ , M , q , and the lag simultaneously (model 1). To achieve a better understanding of the data, we also estimate parameters for a range of fixed values of M (model 3).

This is common practice in the fisheries literature (e.g. Sullivan, 1992). Estimates of q for corresponding values of M can be useful in some contexts where the outcome of an analysis is insensitive to the joint pairs (M, q) (Somers and Wang, 1996). Taking the rough values of Somers and Wang (1996) and Wang and Die (1996) as a guide, we choose the values $M=1, 2$, and 3yr^{-1} . The utility of considering a range of values of M applies equally to considering a range of values for (k, l_{∞}) . Somers and Kirkwood (1991), Wang et al. (1995) and Wang (1998) have all reported estimates of (k, l_{∞}) for this species, and we would like to incorporate this information. However, it is well known that estimates of the growth parameters are strongly correlated. We therefore considered a range of feasible pairs (k, l_{∞}) , and estimated the remaining parameters under model 2. The fixed values we used were, for males, (2, 39.3), (3, 37.7), and (4, 36.1), and for females, (2, 53.1), (3, 47.4), and (4, 41.7). These values were obtained by a

Table 2

Parameter estimates for tiger prawn (*Penaeus semisulcatus*) data. F_{89} is the estimated fishing mortality in 1989. $\text{cor}(M, F_{89})$ is the jackknifed correlation between M and F_{89} . The last column is the objective value per unit of effort. With model 1 all parameters are assumed to be unknown; with model 2 (k, l_{∞}) are assumed to be known; with model 3 M is assumed to be known.

Model	M	F_{89}	Z	k	l_{∞}	σ^*	$\text{cor}(M, F_{89})$	-2log
Males: quasiperiodic recruitment								
1	4.1	2.3	5.2	9.3	33.4	4.5	-0.82	72.96
2	2.9	0.3	3.1	2	39.3	5.1	-0.78	74.43
2	3.7	0.6	3.9	3	37.7	4.3	-0.35	73.99
2	3.4	2.1	4.4	4	36.1	4.3	-0.25	73.60
3	1	2.2	2.0	5.3	32.3	4.8	—	73.05
3	2	1.9	2.9	6.7	32.5	4.8	—	73.03
3	3	0.0	3.0	7.6	32.3	4.8	—	73.15
Males: aperiodic recruitment								
1	1.3	1.6	2.0	5.0	32.6	4.8	-0.67	72.91
2	2.8	0.4	3.0	2	39.3	5.8	-0.79	74.65
2	3.7	0.1	3.7	3	37.7	4.7	-0.81	74.31
2	3.5	0.5	3.7	4	36.1	4.5	-0.64	73.84
3	1	1.8	1.8	4.9	32.4	4.8	—	72.93
3	2	1.0	2.5	5.9	32.5	4.8	—	72.93
3	3	0.0	3.0	7.0	32.5	4.8	—	73.01
Females: quasiperiodic recruitment								
1	4.2	1.7	5.0	5.6	42.2	7.1	-0.65	86.83
2	3.9	0.3	4.1	2	53.1	8.3	-0.83	87.94
2	4.0	0.7	4.3	3	47.4	6.9	-0.66	87.31
2	2.7	1.3	3.3	4	41.7	7.7	-0.71	86.91
3	1	2.6	2.2	4.1	38.8	8.3	—	87.10
3	2	1.8	2.8	4.4	39.9	7.9	—	86.92
3	3	1.6	3.7	5.1	40.7	7.5	—	86.87
Females: aperiodic recruitment								
1	2.6	0.9	3.0	4.9	39.2	8.2	-0.67	86.90
2	3.9	0.1	4.0	2	53.1	13.5	-0.80	88.43
2	4.3	0.1	4.4	3	47.4	9.7	-0.70	87.88
2	3.0	0.8	3.4	4	41.7	8.1	-0.71	87.04
3	1	2.1	2.0	3.6	39.3	8.4	—	87.07
3	2	0.8	2.4	2.8	41.5	8.5	—	86.94
3	3	0.6	3.3	5.0	39.4	8.1	—	86.91

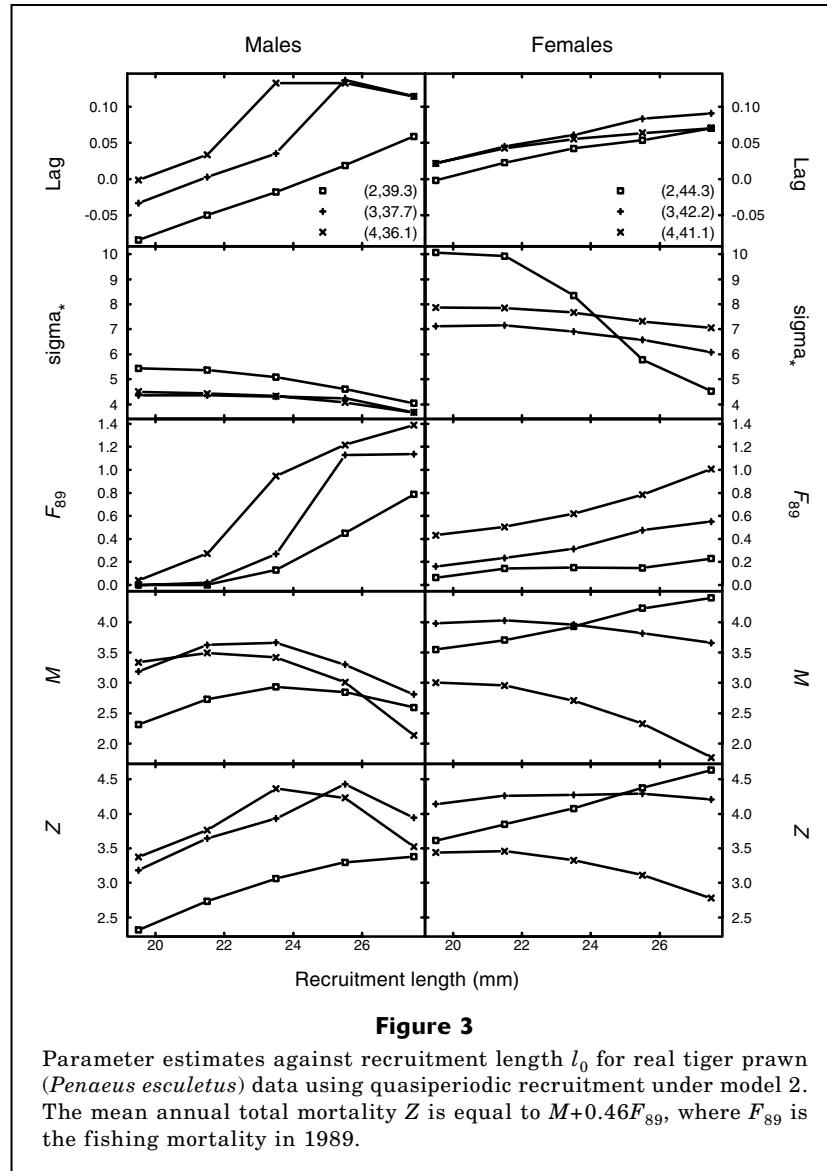
simple linear fit to the estimates in the three papers mentioned above.

The estimates of M from model 1 appear more reasonable for the aperiodic recruitment pattern. The correlations between M and F_{89} (the fishing mortality in 1989) are not as strong as in the simulation example. This is encouraging and indicates that there may be enough contrast in the effort pattern to separate fishing mortality from natural mortality.

The estimates of l_{∞} and σ_* for model 3 are not sensitive to M . However k and M are quite strongly related. In the case of constant recruitment $r(t)$ and mortality $z(t)=Z$, it is well known that k and Z are perfectly correlated, and only their ratio Z/k is able to be estimated. The separation of M and k therefore relies on there being adequate contrast in recruitment and effort.

For model 2 there is little difference between the two recruitment models. The estimates of M show moderate dependence on (k, l_{∞}) , but without trend. These estimates are generally somewhat higher than we expect from prior studies. But for the natural mortality rate, this is the first time we have obtained estimates of M , which is larger than what we have assumed in previous stock assessments, around 2.3 per year (Wang and Die, 1996). Estimates of F_{89} are too variable to be relied upon. All models agree reasonably on the σ_* parameter.

Our model assumes recruitment at a fixed length, l_0 , which has to be chosen. In Figure 3 the parameter estimates for fixed (k, l_{∞}) are plotted against l_0 for the quasiperiodic recruitment model. Parameter estimates are consistent for given l_0 provided that all model assumptions are satisfied. However, when l_0 is too small



or too large, there is bound to be a violation of those assumptions, leading to high sensitivity of the estimates to changes in l_0 . Therefore, we say the most reasonable value for l_0 is that for which the estimates are most slowly varying in the immediate neighborhood of l_0 . On the basis of σ_* , M and Z for males, $l_0=23.5$ would be a reasonable choice. We exclude q from consideration because its standard deviation is comparable to its magnitude (see Table 2). In addition we exclude the lag because we expect it to increase approximately monotonically with l_0 , as indeed it does. There is no clear choice for females; therefore we choose $l_0=23.5$, the same as for males. This choice is consistent with the consideration that l_0 should be somewhere between 20 mm and 30 mm, but in the lower half of the range so that more data can be included in the estimation (because lengths must exceed l_0).

Also shown in Table 2 are jackknife estimates of the standard deviations. The jackknifing is done by dropping the length-frequency record from each occasion in turn and re-estimating the parameters. From the over-all estimate $\hat{\theta}$ and the jackknife estimate $\hat{\theta}_i$ from dropping the i^{th} occasion we obtain a pseudo-value $\hat{\theta} - (n-1)\hat{\theta}_i/n$, where in our case $n=69$. The jackknifed standard deviation is simply the standard deviation of these pseudo-values. We also show the jackknifed correlation between M and q , which is simply the correlation between the corresponding pseudo-values. In most cases there is a large negative correlation.

The fishing mortality in 1989 (the year of peak effort), F_{89} , is simply proportional to q with constant of proportionality 2865, the number of boat-days of effort in that year. The mean total annual mortality Z is $M+0.46F_{89}$ because the mean annual effort was

1320 boat-days. The mostly high negative correlations between M and F_{89} (equivalently, q) may explain why Z tends to have a smaller standard deviation than either M or F_{89} . The results of Figure 3 can be regarded as a sensitivity study on the effect of changing l_0 . The purpose of this sensitivity study is not to estimate l_0 but rather to check that the model assumptions have not been violated for the given l_0 .

The results are fairly similar for the two recruitment models although there are differences: the quasiperiodic recruitment model gives larger F_{89} estimates and smaller σ_s estimates. Our method assumes that the recruitment pattern is known without error; therefore the preferred recruitment pattern should be the one with less error. Let us suppose that the true recruitment pattern consists of a periodic pattern with random variation both within years and between years. If the within-year variation is sufficiently large in comparison with the between-year variation, then the quasiperiodic pattern should be used. On the other hand, if the between-year variation is large, then the aperiodic pattern is preferred. Based on the objective values ($-2\log$) in Table 2, model 2 with quasiperiodic recruitment pattern and fixed k at 4yr^{-1} appears to be the best model for both males and females.

Figure 4 shows the 40 length-frequency records for females with the largest total catch. Overlaid is the expected catch (given the total catch) from the model with (k, l_{∞}) fixed at $(3, 47.4)$ for quasiperiodic recruitment (solid line) and for aperiodic recruitment (dashed line). Because the integral for the expected length distribution is singular in the neighbourhood of l_0 , the first few size classes are omitted from the estimation; only data with length above l_0+2 are used in the estimation. The fit is quite reasonable for most records. It is interesting to compare the performance of the two recruitment models. In early 1988, when recruitment occurred later than usual (see Fig. 2), the aperiodic model tracks the data more closely than the quasiperiodic model, especially in March. On the other hand, the quasiperiodic model fits better in October 1990, whereas the aperiodic model predicts higher abundance of small females because of a recruitment "blip" in September, which was perhaps due to sampling variation.

Discussion

Methods such as McDonald and Pitcher's (1979), ELEFAN (Pauly et al., 1981), and Sparre's (1987) operate on multiple length-frequency data and attempt to identify cohorts in the frequency pattern. Essentially they estimate the growth parameters by tracing cohorts in time; then they estimate mortality by measuring the evolution in abundance of a cohort. For mortality estimation these methods need catch-per-unit-of-effort data. Sparre's method bears some similarity to ours because it attempts to fit the length distribution of a cohort to a normal distribution whose variance is a parameter to be estimated. Our method does not require separation

of cohorts because samples are assumed to come from a length distribution which may be multimodal. Another advantage of our method is that it is not necessary to have information about sampling effort and thus may greatly reduce the complexity of sampling. However, our approach needs a known recruitment pattern.

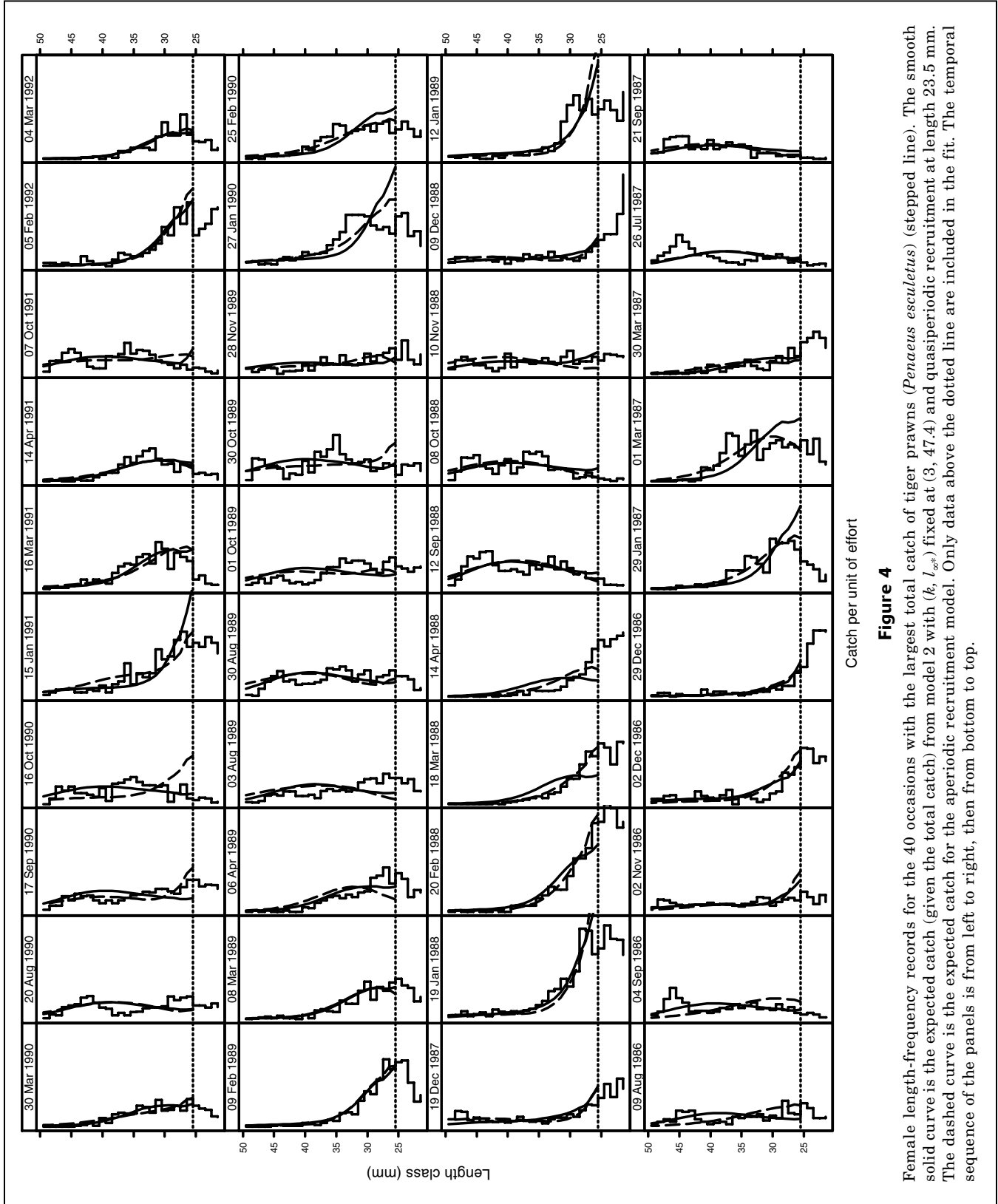
In our application, recruitment was assumed to occur at a fixed length, l_0 , which had to be chosen. We used prior information to constrain l_0 to lie somewhere between 20 mm and 30 mm. We then found the sensitivity of the estimates to changes in l_0 and chose a value that reduced this sensitivity. This choice could be further refined if more accurate constraints were available from other sources. Alternatively, Wang and Somers (1996), who also used l_0 to account for continuous recruitment in estimating growth parameters, have provided guidelines for choosing l_0 .

Deriso and Parma (1988) and Sullivan et al. (1990) reported methods based on stochastic growth. Sullivan (1992) also applied the Kalman filter approach for estimating population parameters. Their models differ from ours in the way random variation is incorporated in the growth model. In their models the length increment from one time step to the next follows a distribution whose mean is given by a fixed growth model. As Wang and Thomas (1995) have demonstrated, this is equivalent to assuming that the growth rate changes randomly from time to time. In our model each individual follows a deterministic growth curve whose L_{∞} parameter is chosen from a random distribution. An individual with larger than average growth at one time step will have above-average growth at subsequent time steps. Perhaps further modeling effort could be directed into combining these approaches.

DeLong et al. (2001) have reported a method for estimating density-dependent natural mortality and the growth rate from length-frequency data for juvenile winter flounder not subject to fishing mortality. Other growth parameters (l_{∞} and the variability of k) were fixed by using information from other sources. Because their data were recorded in the latter half of the year, when recruitment was nearly complete, recruitment was not a complicated issue. In contrast, we had the challenge of a species that recruits all year round. The degree of fit in DeLong et al.'s Figure 5 is comparable to that in our Figure 4.

Our methods are based on distributional assumptions that must be tested for robustness, because, in practice, the l_{∞} distribution of real prawn populations will not equal any of our mathematical distributions. We have found that, even for our ideal model, akin to any other existing model, biases occur for moderate to large coefficients of variation when violation of distributional assumptions occurs.

Our model is motivated by the trawl data from the tiger prawn fishery and relies on 1) known recruitment pattern, 2) contrast in commercial fishing effort for estimation of M and F simultaneously, and 3) contrast in sampling times. Requirement 3 is to spread sampling effort so that growth and mortality information



Female length-frequency records for the 40 occasions with the largest total catch of tiger prawns (*Penaeus esculentus*) (stepped line). The smooth solid curve is the expected catch (given the total catch) from model 2 with (k, l_{∞}) fixed at (3, 47.4) and quasi-periodic recruitment at length 23.5 mm. The dashed curve is the expected catch for the aperiodic recruitment model. Only data above the dotted line are included in the fit. The temporal sequence of the panels is from left to right, then from bottom to top.

are in the data. We fitted a variety of different models. The objective function $-2\log(\text{likelihood})$ values in Table 2 should be used only as guidelines and should not drive the analysis or be used for model selection. Tiger prawns are subject to very high total mortality and hence are short-lived species. Our method is also applicable to longer-lived species. However, for application to other fisheries, some modification of the model may be necessary to incorporate relevant information in the model. Simulation studies may have to be carried out to see how reliable the modified version is for parameter estimation because many factors, such as growth rate and commercial effort patterns, will determine if parameter estimates can be found or how reliable they are if they can be found.

We aim to obtain growth and mortality parameter estimates simultaneously. However, this may be too ambitious, especially for short-lived species unless other information can be incorporated to assist estimation. For instance, Ebert (1973) found estimation of even two parameters (natural and fishing mortality) unreliable and had to assume one of them. This is perhaps why natural mortality is assumed to be known in traditional cohort analysis. Also Askland's method (1994), one of the most recent cohort-analysis methods, requires a known M . Nevertheless, in practice, (k, l_∞) may be estimated from different types of data. The results based on model 2 (assuming (k, l_∞) are known) indicate that both M and F can then be estimated more reliably when there is substantial contrast in the effort pattern. Another assumption is that catchability does not change over time. This may not be necessarily true when new technology is introduced into the fishery (Bishop et al., 2000). The assumption that growth parameters are known greatly reduces the complexity of estimating the remaining unknown parameters and improves the performance of the proposed methods.

We have chosen to allow only l_∞ to be random because, unlike tag-recapture data, the length-frequency data do not have multiple measures from each individual. Each individual is measured only once. Therefore, it might be problematic to allow random K and correlation between K and L_∞ . Such an attempt using length-frequency data may lead to misleading conclusions because the conclusion will be model-driven instead of data-driven. Parameter estimates obtained by fixing M as a constant are deemed more reliable.

We provided a framework for length-frequency data analysis that incorporates continuous recruitment, selectivity, and time-dependent fishing mortality. We have also provided guidelines for how to compute the likelihood function, which depends on rather delicate integrals. Such a model would be very useful for many fisheries because such unified models are not available in the literature. Our work provides a sensible case study. Application of our method may require incorporation of specific information in a fishery. We believe our model, which generalizes the traditional model and is somewhat complicated, has provided us with some use-

ful results for future stock assessment and evaluation of management strategies.

Acknowledgments

This research project was partly supported by the Fisheries Research and Development Corporation of Australia. We gratefully acknowledge the helpful suggestions and comments of David Die, André Punt, Neil Loneragan, and two anonymous referees.

Literature cited

- Askland, M.
1994. A general cohort analysis method. *Biometrics* 50:917-932.
- Beverton, R. J. H., and S. J. Holt.
1956. A review of methods for estimating mortality rates of fish in exploited fish populations, with special reference to sources of bias in catch sampling. *Rapp. P.-V. Réun., Cons. Int. Explor. Mer* 140: 67-83.
- Bishop, J., D. Die, and Y.-G. Wang.
2000. A generalized estimating equations approach for analysis of fishing power in the Australian Northern Prawn Fishery. *Aust. N. Z. J. Stat.* 42:159-177.
- DeLong, A. K., J. S. Collie, C. J. Meise, and J. C. Powell.
2001. Estimating growth and mortality of juvenile winter flounder, *Pseudopleuronectes americanus*, with a length-based model. *Can. J. Fish. Aquat. Sci.* 58:2233-2246.
- Deriso, R. B., and A. M. Parma.
1988. Dynamics of age and size for a stochastic population model. *Can. J. Fish. Aquat. Sci.* 45:1054-1068.
- Ebert, T. A.
1973. Estimating growth and mortality rates from size data. *Oecologia* 11:281-298.
1980. Estimating mortality from growth parameters and a size distribution when recruitment is periodic. *Limnol. Oceanogr.* 26:764-769.
- Hoenig, J. M.
1987. Estimation of growth and mortality parameters for use in length-structure stock production models. *In* Length-based methods in fisheries research (D. Pauly and G. R. Morgan, eds.), p. 121-128. ICLARM Conf. Proc. 13.
- McDonald, P. D., and T. J. Pitcher.
1979. Age groups from size-frequency data: a versatile and efficient method of analyzing distribution mixtures. *J. Fish. Res. Board Can.* 36:987-1001.
- Pauly, D., J. Ingles, and T. Neal.
1981. Application to shrimp stocks of objective methods for the estimation of growth, mortality and recruitment-related parameters from length-frequency data (ELEFAN I and II). *In* Penaeid shrimps: their biology and management (J. A. Gulland and B. J. Rothschild, eds.), p.220-234. Fishing News Books Ltd., Farnham, England.
- Powell, D. G.
1979. Estimation of mortality and growth parameters from the length frequency of a catch. *Rapp. P.-V. Réun., Cons. Int. Explor. Mer* 175:167-169.

- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery.
1992. Numerical recipes in C; the art of scientific computing (2nd ed.). Cambridge Univ. Press, Cambridge, England.
- Sainsbury, K. J.
1980. Effect of individual variability on the von Bertalanffy growth equation. *Can. J. Fish. Aquat. Sci.* 37:241-247.
- Somers I. F., and G. P. Kirkwood.
1991. Population ecology of the grooved tiger prawn, *P. semisulcatus*, in the North-western Gulf of Carpentaria, Australia: growth, movement, age structure and infestation by the bobyrid parasite *Epipenaeon ingens*. *Aust. J. Mar. Freshw. Res.* 42:349-367.
- Somers, I. F., and Y.-G. Wang.
1996. A bioeconomic analysis of seasonal closures in Australia's multispecies Northern Prawn Fishery. *N. Am. J. Fish. Manag.* 17:114-130.
- Sparre, P.
1987. A method for estimating growth, mortality and gear selection/recruitment parameters from length-frequency samples weighted by catch per effort. *In* Length-based methods in fisheries research. ICLARM Conf. Proc. 13 (D. Pauly and G. R. Morgan, eds.), p. 75-102.
- Ssentongo, G. W., and P. A. Larkin.
1973. Some simple methods of estimating mortality rates of exploited fish populations. *J. Fish. Res. Board Can.* 30:695-698.
- Sullivan, P. J.
1992. A Kalman filter approach to catch-length analysis. *Biometrics* 48:237-257.
- Sullivan, P. J, H. L. Lai, and V. F. Gallucci.
1990. A catch-at-length analysis that incorporates a stochastic model of growth. *Can. J. Fish. Aquat. Sci.* 47:184-198.
- Vetter, E. F.
1988. Estimation of natural mortality in fish stocks: a review. *Fish. Bull.* 86:25-43.
- Wang, Y.-G.
1998. An improved Fabens method for estimation of growth parameters in the von Bertalanffy model with individual asymptotes. *Can. J. Fish. Aquat. Sci.* 55:397-400.
- Wang, Y.-G., and D. Die.
1996. Stock-recruitment relationship of the tiger prawns (*Penaeus esculentus* and *Penaeus semisulcatus*) in the Australian Northern Prawn Fishery. *Aust. J. Mar. Freshw. Res.* 47:87-95.
- Wang, Y.-G., and N. Ellis.
1998. Effect of individual variability on estimation of population parameters from length-frequency data. *Can. J. Fish. Aquat. Sci.* 55:2393-2401.
- Wang, Y.-G., and I. F. Somers.
1996. A simple method for estimating growth parameters from multiple length-frequency data in presence of continuous recruitment. *Fish. Res.* 28:45-56.
- Wang, Y.-G., and M. R. Thomas.
1995. Accounting for the individual variability in the von Bertalanffy growth model. *Can. J. Fish. Aquat. Sci.* 52:1368-1375.
- Wang, Y.-G., M. R. Thomas, and I. F. Somers.
1995. A maximum likelihood approach for estimating growth parameters from tag-recapture data. *Can. J. Fish. Aquat. Sci.* 52:252-259.
- Wetherall, J. A., J. J. Polovina, and S. Ralston.
1987. Estimating growth and mortality in steady-state fish stocks from length-frequency data. *In* Length-based methods in fisheries research (D. Pauly, and G. R. Morgan, eds.), p. 53-75. ICLARM Conf. Proc. 13.