Use of length-frequency analysis for estimating the age structure of the catch of *Nephrops norvegicus* (Crustacea: Nephropidae)

Margarida Castro


The capacity of length-frequency analysis techniques to provide acceptable estimates of the age structure of the catch of *Nephrops norvegicus* (Crustacea: Nephropidae) was evaluated. Special attention was given to the stock from the south coast of Portugal, designated as the Algarve stock. Samples representing the length frequencies from the catch of a *Nephrops* stock with known age structure were generated. The simulations incorporated all the biological and fishing information available for the stock of interest. A selection of length-frequency analysis techniques was applied to the simulated samples: the Bhattacharya method of identification of the components of a mixture; MIX, a computer program to perform the same task; and the MULTIFAN program. The methods were used on 20 replicates of a 12-month series of length-frequency samples. Of the methods used, MIX produced the best results. The analysis of simulated data suggests that, if the assumptions for the simulations are correct, a mixture separator can be used to estimate the age structure of *Nephrops* populations. Nevertheless, length-frequency analysis should not be utilized without information on the biology of the species, such as intermolt duration and increment at molt.

Margarida Castro: Graduate School of Oceanography, Narragansett Bay Campus, Narragansett, RI 02882, USA. Present address: UCTRA, Universidade do Algarve, 8000 Faro, Portugal [tel: (+351) 89 800 970, fax: (+351) 89 818 353].

Introduction

*Nephrops norvegicus* (common name Norway lobster) is an important species from a commercial point of view, as shown by the large amount of work that has been dedicated to this species, particularly within the ICES Working Group on *Nephrops* and *Pandalus* Stocks. Despite being one of the most studied decapods, the area of age and growth estimation is still one for which there is no standard methodology. The difficulties with this area of work are intrinsic to crustaceans. Since they do not have permanent hard structures, age estimation by observation of growth rings on calcified parts is not possible. The importance of growth studies has led many authors to try alternative techniques to direct age estimation. Three methods of analysis have been used with *Nephrops norvegicus*: (1) the analysis of length-frequency data (Hillis, 1971 and 1972b; Farmer, 1973; Charuau, 1975; Conan, 1978a; Nicholson, 1979; Figueiredo, 1984); (2) direct observation of animals in captivity (Thomas, 1965; Hillis, 1971, 1972a, 1973, 1974; Figueiredo, 1975; Charuau, 1977; Charuau and Conan, 1977; Sarda, 1980); and (3) mark-recapture studies (Andersen, 1962; Hillis, 1974, 1979; Charuau and Conan, 1977). The last two methods have some disadvantages. Molting in captivity may not be representative of molting in the natural environment and tagging may seriously impair growth rates and even the ability to molt. Thus, the use of length-frequency data analysis remains an important tool for solving this type of problem. The main objective of this work was to evaluate the capacity of length-frequency analysis techniques to provide acceptable estimates of the age structure of the catch of *Nephrops norvegicus*, with particular emphasis on the population off the south of Portugal, designated as the Algarve stock.

The methodology consisted of the following four steps: (1) Generation of simulated samples representing a series of length frequencies from the catch of a *Nephrops* stock, with known age structure; (2) application of selected length-frequency analysis methods to the simulated samples; (3) comparison of observed and estimated age structures to evaluate the methods in terms of parameter recovery; (4) discussion of the suitability of...
length-frequency analysis methods to estimate the structure of the catch, with particular attention to the Algarve stock.

Materials and methods

Length frequencies from the commercial landings

Part of this work involved the analyses of length frequencies from landings of *Nephrops norvegicus* obtained in the port of Olhao, Portugal, from January 1981 to December 1985. The lobsters were caught by bottom trawlers that fish the stock off the Algarve coast. The samples were collected as part of a program called the “National Plan for Biological Sampling” (Plano Nacional de Amostragem Biológica – PNAB), a responsibility of the National Institute for Fisheries Research (Instituto Nacional de Investigação das Pescas – INIP), in Lisbon. The length structure of the samples was extrapolated to the total catch landed on the south coast of Portugal. These data were used to provide information used in the simulations relative to sex ratio of the catch and distribution of the fishing effort throughout the year.

Simulated length frequencies

Simulated length frequencies were generated for a 12-month period. The simulated samples represent the length frequency of the catch based on the biological and fisheries parameters considered most probable for the Algarve stock. The simulated data consisted of 20 independent time series, each one representing the catch for 12 consecutive months. Males and females were analyzed separately.

In order to obtain information for the simulations, data from the sampled landings were used preferentially. When the necessary information could not be obtained in this way, published data for *Nephrops norvegicus* were used. In areas that involve the estimation of life history parameters, a literature search was conducted to compile as many estimates as possible. Whenever estimates for the stock off southern Portugal were available, they were preferred. Behavioral aspects were assumed not to differ from other populations that have been studied.

The choice of parameters was at times arbitrary (as for example the increment at molt model), but was always influenced by the guidelines defined in the previous paragraph. The estimation of parameters for the different aspects of the simulation was based on the following: (1) Female life cycle (distribution of size of first maturity) – research cruise data from Figueiredo (1982) for the Algarve stock; (2) intermolt period – lobsters in captivity, data from Sardà (1983); (3) increment at molt – estimated on the basis of the length structure of the catch analyzed in this work, taking into consideration the intermolt period as defined before; (4) size at settlement on the bottom – lobsters in captivity, data from Figueiredo (1975); (5) natural mortality – published data for *Nephrops norvegicus* and the possible range of age classes in the catch; (6) fishing mortality – estimated for the Algarve stock by Caramelo (1986).

The approach used in this work consisted of testing length-frequency analysis techniques. The simulations included biological and fishery parameters, based on a review of the *Nephrops* literature and on data collected for the Algarve stock. Most of the processes – and, in particular, growth – included in the simulations were treated in a probabilistic manner. Growth in *Nephrops* was broken into two components, intermolt period and increment at molt. Each one of these factors was modeled separately, using the model to predict mean values, and allowing for random variability around the mean. This approach was proposed by Caddy (1987) and was applied to Florida male stone crabs (Restrepo, 1989).

Other aspects included in the simulations were distribution of length of first maturity, probability of completion of the maturation cycle for females, probability of death due to natural causes or due to fishing, distribution of age at capture, and distribution of initial recruitment numbers. Details of the assumptions for the simulations are provided by Castro (1990).

The simulated data were not affected by sampling errors. This was a deliberate decision, since it was considered fair to evaluate the length-frequency analysis methods under the best possible conditions, under the assumption that if these methods work it would be worth exploring their applicability to real data. *Nephrops norvegicus* off the south coast of Portugal are known to have one spawning season per year (Figueiredo and Barraca, 1963). This information was taken into consideration in the simulations and can be used to make decisions in length-frequency analysis both with simulated and real data. Therefore, the components (modes) identified in a given length distribution were considered to correspond to distinct age classes. The number of age classes contributing to any of the analyzed length distributions was always seven. When simulating the data, age seven was considered to be the oldest age class present in the catch, and age zero was considered to be not affected by fishing. Therefore, the catch in all cases is composed of age classes one to seven. In some of the winter months, age 6 and/or age 7 were not present in the catch, but in the month of August (chosen for the initial testing of the methods), when the fishing effort was considered to be highest, all seven age classes were consistently present.

For all methods tested, preliminary work was done to decide which class interval width should be used. An
initial analysis of the plots was done for interval sizes of 1, 2, 3, and 5 mm. These cover all the size classes reported to have been used with *Nephrops norvegicus*. Only size classes 1 and 2 mm were tested with the different methods. Data grouped in class sizes 3 and 5 mm showed loss of information and were abandoned.

The growth model used to simulate the data (stochastic growth as a function of intermolt period and increment at molt) may lead to mean length-at-age values that do not fall on a von Bertalanffy growth curve. In such a case, the parameters of the von Bertalanffy growth curve change with the range of age classes chosen for their estimation. This phenomenon has been verified in real life situations for other species (Hirschhorn, 1974).

Most of the methods that estimate parameters of a mixture of age classes in a length distribution work under the assumption that the distribution of length-at-age is normal. To verify this assumption in the simulated data, tests for normality were performed on the simulated lengths of 100 males and 100 females at 1-year intervals. For age classes 1 to 7, tests of significance for skewness and kurtosis were performed (Sokal and Rohlf, 1981). For males, no age class showed significant deviations from normality ($\alpha = 0.05$). Ages 5, 6, and 7 for females showed a slight positive skewness, significant at the 0.05 level, but not significant at the 0.001 level. This is due to the distribution of length of first maturity, considered to be positively skewed. When this distribution of age of first maturity was used in the data simulation, individuals that mature later grew with a higher growth rate for longer, and appeared as rare large individuals in adult age classes.

**Methodology**

Keeping in mind that the objective of this work was to select length-frequency analysis methods that can be used to estimate the age structure of the catch of a *Nephrops* population, the following methodology was used: (1) Use of selected methods on the simulated data. The number of monthly samples simulated in this work was 480 (12 months $\times$ 20 replicates $\times$ 2 sexes). For methods that work with one sample at a time, analyzing all the samples would require a great deal of time. Therefore, the methods were tested on one month for each case and sex. The month of August for the first year of simulated data was used, as August is the month in which the catch is generally the largest and females are well represented; (2) evaluation of the "performance" of each method by comparing the predicted age structure obtained with the known age structure of the catch.

Throughout this work the term length refers to the standard length for this species, namely the carapace length.

As discussed previously, the simulated data were not adequately described by the von Bertalanffy growth curve, therefore length-frequency analysis methods that only provide estimates of the von Bertalanffy parameters were not considered. Only methods that could be used to estimate the age structure of the catch at given points in time were selected. These were:

1. Bhattacharya's (1967) method, representing the graphical methods. Compared with the other two graphical methods frequently used in fisheries, fitting of parabolas to log-transformed distributions (Tanaka, 1956) and the probability paper method (Harding, 1949; Cassie, 1950, 1954), Bhattacharya's approach is simpler. Also, an objective criterion can be developed easily for the application of this method, as shown by Pauly and Caddy (1985). The Bhattacharya method is also widely used, and has been incorporated in two major computer packages developed for stock assessment in fisheries, LFSA (Sparre, 1987) and the Compleat ELEFAN (Gayamilo *et al.*, 1988).

2. MIX, representing the methods based on maximum likelihood estimation. All numerical analysis methods use one of two estimation techniques, maximum likelihood estimation or least-squares estimation. Least-squares estimation was compared with maximum likelihood estimation by Akamine (1985) and was found to produce worse results. Therefore it was not considered. The maximum likelihood methods are all similar, differing mainly in the optimization technique utilized. MIX was considered to represent all the approaches that are based on maximum likelihood estimation. The variation of this technique introduced by Schnute and Fournier (1980), with constraints on means fit to the von Bertalanffy growth model, was not tested because the von Bertalanffy growth curve did not describe the growth of the simulated data well.

3. MULTIFAN (Otter Software, 1988). MULTIFAN also uses maximum likelihood estimation, but it works on a series of length-frequency samples and uses the information in all the samples for estimation of the parameters.

**Application of methods to data**

Bhattacharya's method was used only with data for the month of August. Two size intervals, 1 mm and 2 mm, were tested initially. With 2 mm classes, a series of three points only rarely fell on a straight line with a negative slope. An approach similar to what was suggested by Pauly and Caddy (1985) was followed for the analysis of the data. It consisted of statistically testing the correlation coefficient of all possible lines with negative slopes.
made up of 3–7 consecutive points. The critical level chosen was \( a = 0.05 \). The program BHATTREG.BAS (Erzini, 1990) was used for the analysis. The accepted lines were sorted by the estimated mean value of the component identified. In most cases the same point belonged to more than one of the accepted lines. In that case, where more than one significant regression was identified in the same area of the graph, the line with the highest regression coefficient was used to estimate the mean and standard deviation of that component. The means of the identified components are compared with the observed mean lengths-at-age in Figure 1.

With MIX, again only the month of August for the first year of the simulations was used. The first analysis of samples, using 1 mm length classes, proved to be difficult. There were too many peaks in the samples, and the means were not clear, so the interval width chosen for the analysis was 2 mm. The first step was the visual analysis of graphs of the length distributions. Initial estimates of the number of components, and their means and standard deviations were made. In estimating the number of components in each sample, the known number of ages (seven) was ignored. Estimation of the number of components was based solely on visual analysis of the sample. The initial estimates of the number of components varied from 6 to 9. All peaks and deviations from normality were assumed to indicate the presence of an age group. This was done in order to represent a situation where information about the age structure of the population is not known. Whenever the program produced a negative value for the proportion attributed to one of the components, that component was removed and the parameters re-entered from the beginning. The

Figure 1. Mean length-at-age observed and estimated using the Bhattacharyya method. Only results for the month of August are presented.

Figure 2. Mean length-at-age observed and estimated using MIX. Only results for the month of August are presented.
second step consisted of the analysis of the samples using MIX. The selection of routines for the estimation of parameters was based on the suggested steps included in the program's manual (MacDonald and Green, 1985). A comparison of observed and estimated mean lengths-at-age in each sample is presented in Figure 2.

MULTIFAN was run on all the 20 series of 12 months of simulated data. The program allows for a great deal of flexibility in the choice of parameters and their variability. The options chosen in this analysis that are different from the basic defaults were: (1) The potential number of age classes considered was 6 to 7 for females and 7 to 8 for males. These numbers were selected by visual analysis of plots of the simulated length frequencies, the same methodology used with MIX. Initial estimates for K were chosen to cover a range of K values from 0.01 to 0.3 in males and 0.005 to 0.5 in females; (2) the predicted mean lengths-at-age for all age classes were allowed to deviate from the von Bertalanffy growth model with no penalty for the computation of the likelihood function value (switch 9 = 8 and switch 7 = 0). No other constraints were imposed on mean length-at-age; (3) all length frequencies were included in the estimation, and the accuracy of each sample was assumed to be proportional to the sample size (switch 30 = 1); (4) standard deviation of length-at-age was assumed to be length-dependent (switch 20 = 5). No other constraints were imposed on the standard deviation of mean length-at-age; (5) for females, seasonal variability of the growth rate was allowed (switch 21 = 1 and switch 22 = 1). The program was run on data grouped into 1 and 2 mm class sizes. Since no significant differences were observed, and the running of the program with 2 mm length classes was much faster, 2 mm length classes were used. The estimated and observed mean length-at-age values for the month of August are presented in Figure 3.

Results

The Bhattacharya method may be unfairly evaluated by setting such a rigid criterion, but this was the only way of avoiding subjectivity in the analysis. If the method had been applied in the traditional way, by visual selection of series of points from the length frequencies, it is possible that a larger number of length classes would have been identified. It should be stressed that, in spite of the fact that few age classes were identified, those that were identified were done correctly (Fig. 1).

MIX clearly produced the best results. Age classes up to 3 in females and 4 in males were consistently identified correctly. It should be stressed that these age classes include approximately 92% of the catch for males and 73% of the catch for females. The modes in older classes are still fairly well separated, in particular for males. The inability of the techniques used in MIX to identify older age classes has to do with their small numbers combined with greater variability of length-at-age (Fig. 2).

MULTIFAN also showed some difficulties. The results presented in Figure 3 refer only to the month of August, but the mean lengths-at-age were estimated taking into consideration all the other samples in the series (months January to December) and assuming that the evolution of mean length-at-age can be represented by a von Bertalanffy growth curve. The release of the mean lengths-at-age from the von Bertalanffy growth model was not enough to improve the estimation of mean length-at-age.

MIX was chosen for estimation of the age structure of the catch for all simulated samples. Since MIX has to be applied to every monthly distribution, if all the samples were analyzed, 480 runs would be required (2 sexes × 12
Figure 4. Mean length-at-age for males, observed and estimated, for a series of 12 months of simulated length-frequency distributions. The estimated values were obtained using MIX.
Figure 5. Mean length-at-age for females, observed and estimated, for a series of 12 months of simulated length-frequency distributions. The estimated values were obtained using MIX.
months × 20 cases). To reduce the amount of time required to perform the analysis in all samples, only the first five cases for each sex were analyzed. The results are presented in Figures 4 and 5. Cohorts corresponding to ages 1 to 3 in females and 1 to 4 in males can clearly be followed in the samples which confirm the results obtained for the single month analyses. In addition to providing information on mean length-at-age, MIX also produces estimates of standard deviation of length-at-age, proportions for each age class, as well as standard errors for all the estimated parameters.

Discussion

Assuming that the biological parameters used in the simulations are realistic, it is possible to suggest that a distribution mixture separation program like MIX can be used to estimate the age structure of the catch, if good samples are collected and additional information on growth exists. This additional information can be used to guide in the choice of initial estimates and to set boundaries on the parameters being estimated.

The two components of growth, increment at molt and intermolt period can be studied by direct observation of wild animals. The increment at molt can be estimated by keeping individuals in captivity, as done in several studies mentioned previously. To minimize the changes in growth due to captivity, only data from the first molt in captivity, and for animals that molted within a short period after capture should be used. If the new carapace is already completely formed at the point of capture, it seems reasonable to assume that the postmolt length would not be significantly different from the value had the animal molted under natural conditions. Such an experimental setup would also provide information on the variability of increment at molt, necessary for estimation of the variability of length-at-age.

The second aspect of growth, the duration of intermolt period, can be estimated by direct observation of wild animals. The count of soft animals does not seem to be a reasonable way of detecting molt frequency, since the period of time the animals stay soft is short, and a sampling design to detect such a relatively rare situation would be difficult to implement. As an alternative, the intermolt period can be studied by observation of the pleopods at regular time intervals using the technique proposed by Drach (1939) to determine the intermolt stages. This approach has been used successfully for Nephrops (Hillis, 1971; Charuau, 1973; Sardà, 1985).

In conclusion, it is suggested that length-frequency analysis can be used to estimate the age structure of the catch of Nephrops with a biological cycle similar to the one of the Algarve stock. Even if the age class separation is incomplete, the youngest age classes that contribute most to the catch can be identified. It should be noted that the lobsters off the Portuguese coast, in particular those from the west stock, reach very large sizes. Females with more than 70 mm carapace length and males larger than 90 mm have been measured. Such large sizes have never been reported for populations from the northern latitudes of this species. In this situation we are dealing with either higher growth rates or greater longevity. Higher growth rates would contribute to improving the chances of better results with length-frequency analysis, because the mean lengths-at-age would be further apart, facilitating the identification of age classes in the length distributions. Nicholson (1979) did a theoretical study of the use of length-frequency analysis for age estimation of Nephrops. Using simulated data based on a theoretical growth model which utilized published information from several sources relative to Nephrops from the northern extremes of this species distribution, Nicholson suggested that it is unlikely modes corresponding to year classes 2+ and older could be detected.

It should be stressed that, although length-frequency analysis seems to be a useful tool for estimating the age structure of Nephrops populations, these techniques should not be used without the support of biological information collected independently. Information on increment at molt and intermolt period duration (mean values and variability) can be used to predict mean length and variability of length-at-age, as well as the number of age classes present in a given area. This information, used in conjunction with a mixture separator such as MIX, could eventually be used to estimate age-length keys for Nephrops.

References


