Fisheries and Oceans Canada / Pêches et Océans Canada

Canadian Stock Assessment Secretariat
Research Document 99/115

Secrétariat canadien pour l'évaluation des stocks
Document de recherche 99/115

Not to be cited without
Permission of the authors[1]

Ne pas citer sans
Autorisation des auteurs[1]

Developing an overall indicator for monitoring temporal variation in fish size at age and its application to cod (*Gadus morhua*) in the Northwest Atlantic, NAFO subdivision 3Ps

Yong Chen and Luiz G. S. Mello

Fisheries Conservation Chair

Fisheries and Marine Institute

Memorial University of Newfoundland

St. John's, Newfoundland

Canada A1C 5R3

[1] This series documents the scientific basis for the evaluation of fisheries resources in Canada. As such, it addresses the issues of the day in the time frames required and the documents it contains are not intended as definitive statements on the subjects addressed but rather as progress reports on ongoing investigations.

Research documents are produced in the official language in which they are provided to the Secretariat.

[1] La présente série documente les bases scientifiques des évaluations des ressources halieutiques du Canada. Elle traite des problèmes courants selon les échéanciers dictés. Les documents qu'elle contient ne doivent pas être considérés comme des énoncés définitifs sur les sujets traités, mais plutôt comme des rapports d'étape sur les études en cours.

Les documents de recherche sont publiés dans la langue officielle utilisée dans le manuscrit envoyé au secrétariat.

ISSN 1480-4883
Ottawa, 1999

Canada

Abstract

Size-at-age data are commonly available life-history information for a fishery. They are essential in formulating management policies for fish stocks. Because fish size-at-age values are likely to be influenced by many biotic and abiotic environmental variables, monitoring temporal variations in size-at-age can often provide fisheries managers with important information about status of fish stocks and help them identify the necessary changes in management policies. However, because of the multivariate nature of size-at-age data, it is often difficult to objectively evaluate temporal variations in size across all age groups. Commonly used single-age-based approaches ignore covariances among sizes of different age groups and may derive different results in evaluating temporal variations when different age groups are used. In this study, we propose a two-step approach to developing an indicator for monitoring temporal variation in fish size-at-age data. A robust approach, minimum volume ellipsoid analysis, is used to identify outliers in size-at-age data, and then a weighted principal component analysis is applied to the data with identified outliers down-weighted. An indicator is defined from the resultant principal components for monitoring temporal variations in size-at-age data. The proposed approach was applied to size-at-age data of cod (*Gadus morhua*) in the Northwest Atlantic, NAFO subdivision 3Ps. An overall size-at-age indicator was developed for the cod stock. Sizes for the pre-1980's year classes were found to be much higher than those for the post-1980 year classes.

Résumé

Les données sur la taille selon l'âge sont des données concernant le cycle vital qui sont généralement disponibles pour les pêches. Elles sont essentielles à la formulation des politiques de gestion des stocks de poisson. Comme la taille selon l'âge des poissons est normalement influencée par bon nombre de variables biotiques et abiotiques, le suivi des variations temporelles de ce paramètre peut souvent constituer une source d'information importante pour les gestionnaires des pêches relativement à l'état des stocks de poisson et les aider à déterminer les modifications devant être apportées aux politiques de gestion. Cependant, parce que les données relatives à la taille selon l'âge étant des valeurs à variables multiples, il est souvent difficile d'évaluer objectivement les variations temporelles de taille au sein de tous les groupes d'âges. Les méthodes généralement utilisées, fondées sur un seul âge, ignorent la covariance entre les tailles de groupes d'âges différents et peuvent donc donner des résultats différents au moment de l'évaluation des variations temporelles si des groupes d'âge différents sont utilisés. Dans la présente étude, nous proposons une méthode en deux étapes pour l'élaboration d'un indicateur de suivi de la variation temporelle des données sur la taille selon l'âge. Une méthode robuste, par analyse de l'ellipsoïde de moindre volume, est utilisée pour identifier les valeurs aberrantes des données de la taille selon l'âge et, ensuite, une analyse par élément principal pondéré est appliquée aux données pour pondérer à la baisse les valeurs aberrantes. Un indicateur est défini à partir des éléments principaux pour le suivi des variations temporelles des données sur la taille selon l'âge. La méthode proposée a été appliquée aux données sur la taille selon l'âge de la morue *(Gadus morhua)* de la sous-division 3Ps de l'OPANO, dans l'Atlantique nord-ouest. Un indicateur général de la taille selon l'âge a été élaboré pour le stock de morue. Les effectifs des classes d'âge d'avant les années 1980 se sont avérés de beaucoup supérieurs à ceux des classes d'après les années 1980.

Introduction

In fisheries, growth can mean change in size and/or change in number. The former is often referred to as change in a specified dimension of fish body over a certain period of time, and the latter is referred to as change in the total number of fish within a fish population or stock over a defined time period. Although different, these two processes are two of the most important factors that result in an increase in the biomass of a fish stock. The focus of this study is, however, related to the growth in size.

As a result of growth, individual fish acquires a certain size (length) at a certain age and the size of a fish increases with age. Such measurements of size are often referred to as size at age (Hilborn and Walters 1992). Obviously, the faster a fish grows, the larger its size-at-age value is. Variations in growth rates can exist even among individuals that are born at the same time, resulting in variability in their size-at-age values. For a fish population, the size of a given age group is often measured as an average of sizes of all individuals in this age group. In practice, it is impossible to measure all fish in an age group. Often a certain number of fish are randomly sampled, their sizes measured and ages determined. The average size for each age group is then calculated. The resultant size-at-age data are used as estimates of size-at-age for the population (Ricker 1975).

Size-at-age data are important in understanding the dynamics of fish population. They are essential in estimating fish stock biomass and productivity. By evaluating differences in sizes of two age groups, growth in size between these two ages can be estimated. This can provide fisheries managers with important information such as fish growth rate and the age at which fish attain their highest growth rate (Nikolskii 1965, Paloheimo and Dickie 1965, Myers et al. 1997). Such information is essential in formulating management policies (Hilborn and Walters 1992). Many fish stock assessment models such as yield-per-recruit models and delay-difference models require size-at-age as input data (Ricker 1975, Hilborn and Walters 1992).

Many biotic and abiotic environmental variables may affect growth in size, and subsequently fish size-at-age values (Nikolskii 1965, Paloheimo and Dickie 1965, Moreau 1987). A direct consequence of such a process is fluctuation in fish size-at-age values among different year classes (Beacham 1983, Chen and Harvey 1995). A close monitoring of temporal changes in size-at-age values may reveal some important information about the status of fish stocks and can help fisheries managers identify the necessary changes in management policies (Beverton and Holt 1957, Ricker 1975, Charnov 1993). For example, a substantial decrease in fish stock biomass may lead to a decrease in the age of fish attaining maturity which may in turn result in a decrease in size at age (Nikolskii 1965, Roff 1984, Chen and Harvey 1994, Jensen 1996).

Because of the multivariate nature, it may be difficulty to evaluate temporal changes in fish size-at-age data. In practice, such an evaluation is often conducted by examining the size of fish at each age separately or evaluating the size of fish in an age group arbitrarily selected by the researcher (Beacham 1983, Lilly 1996). Such a single-age-based approach ignores the relations among sizes of different ages (e.g. size at age 1 affects size at age 2). Different results may arise from the use of different age groups in evaluating temporal patterns.

Fisheries data of multivariate nature are often analyzed using principal component analysis (PCA; Manly 1991, Jackson 1993, Chen and Harvey 1995). This method is one of the most common data-exploratory multivariate ordination techniques allowing study of data relationship and the reduction in dimensionality (Rao 1964, Jackson 1993). Such a multivariate approach can reduce the number of variables, while essential information that is inherent in the original data is kept.

In this study, we propose using PCA to summarize size-at-age data. Because fisheries data tend to be subject to atypical errors (Chen and Harvey 1994, Chen et al. 1994), a two-step procedure is proposed: a robust multivariate approach is applied to size-at-age data to identify outliers in data, and then a weighted PCA is applied with the defined outliers down-weighted. The resultant principal components

(PCs) are interpreted with respect to the original variables (Manly 1991, Jackson 1993). An indicator is identified from the resultant PCs for monitoring temporal variations in size-at-age data. The proposed approach is applied to size-at-age data of cod (*Gadus morhua*) in the Northwest Atlantic, NAFO subdivision 3Ps. An overall size-at-age indicator is developed for the cod stock.

Methods and materials

Identification of outliers for multivariate data

Fisheries data are commonly subject to errors of various sources (Hilborn and Walters 1992, Chen and Paloheimo 1998, Walters 1999). This may result in outliers in modelling fisheries data (Chen et al. 1994). Commonly used statistical methods such as PCA can be severely biased by the existence of outliers in data (Rousseeuw and Leroy 1987). Outliers are much more difficulty to identify in a multivariate analysis than in a univariate analysis (Rousseeuw and Leroy 1987). The commonly used method for identifying outliers in multivariate analyses is the squared Mahalanobis distance (Krzanowski 1988). For a data matrix with K variables and each variable having n observations,

$$\mathbf{X} = \begin{pmatrix} x_{1l} & \cdots & x_{il} & \cdots & x_{Kl} \\ . & . & . & . & . \\ x_{1j} & \cdots & x_{ij} & \cdots & x_{Kj} \\ x_{1n} & \cdots & x_{in} & \cdots & x_{Kn} \end{pmatrix} = \begin{pmatrix} \mathbf{x}_1 \\ . \\ \mathbf{x}_j \\ . \\ . \\ \mathbf{x}_n \end{pmatrix}$$

the squared Mahalanobis distance is calculated as

$$MD^2(\mathbf{x}_j, \mathbf{X}) = (\mathbf{x}_j - \mathrm{T}(\mathbf{X}))\, \mathrm{C}^{-1}(\mathbf{X})\, (\mathbf{x}_j - \mathrm{T}(\mathbf{X}))^t$$

for each point $\mathbf{x}_j$, where bold letters are vectors or matrices, T(**X**) is the arithmetic mean of the data set **X** and **C**(**X**) is the classical covariance estimate. These are calculated as

$$T(\mathbf{X}) = \frac{1}{n}\sum_{j=1}^{n} \mathbf{x}_j$$

$$\mathbf{C}(\mathbf{X}) = \frac{1}{n-1}\sum_{j=1}^{n} (\mathbf{x}_j - \mathrm{T}(\mathbf{X}))^t (\mathbf{x}_j - \mathrm{T}(\mathbf{X})) \ .$$

Points having large $MD^2(\mathbf{x}_j, \mathbf{X})$ value are identified as outliers and are subsequently deleted from the sample for further analyses. This approach works well if only a single outlier is present (Rousseeuw and Zomeren 1990), but perhaps suffers from the masking effect if there are more than one outliers because one far-away outlier can make all other outliers have small $MD^2(\mathbf{x}_j, X)$. Some refinements of this approach have been proposed such as iterative deletion, iterative trimming, and depth trimming (Campbell 1980, Devlin et al. 1981, Rousseeuw 1985). However, problems associated with the MD^2 approach still exist in these methods (Rousseeuw and Zomeren 1990).

A robust method, the minimum volume ellipsoid (MVE), was proposed to identify outliers in estimating means and covariance for multivariate data by Rousseeuw (1984, 1985). Although it is not uncommon that data from fisheries or ecological studies are contaminated by outliers, the effects of outliers on multivariate analysis (e.g. PCA, canonical correspondence analysis, and multiscaling methods) have received little attention. The MVE has recently been used in other research fields such as engineering and economics (Rousseeuw and Leroy 1987), but its application in fisheries or ecological studies is limited.

An algorithm that involves extensive computer subsampling has been suggested for the MVE analysis. This algorithm can be summarized as follows:

(1) For a multivariate data matrix $\mathbf{X}$ with K variables and n observations (as described above), draw a subsample of K+1 different observations, indexed by $J = (j_1, ..., j_{K+1})$, and calculate the arithmetic mean and the corresponding covariance matrix as $T_J = \frac{1}{K+1}\sum_{j\in J}\mathbf{x}_j$ and $\mathbf{C}_J = \frac{1}{K}\sum_{j\in J}(\mathbf{x}_j - \mathrm{T}_J)^t(\mathbf{x}_j - \mathrm{T}_J)$ where $\mathbf{C}_J$ is nonsingular;

(2) Calculate $m_J^2 = [(\mathbf{x}_j - \mathrm{T}_J)\mathbf{C}_J^{-1}(\mathbf{x}_j - \mathrm{T}_J)^t]_{h:n}$
where h = (n + K +1)/2 and the above computation corresponds that the ellipsoid should be inflated or deflated to contain exactly h points (out of n points);

(3) Calculate $P_J = (\det(m_J^2\mathbf{C}_J))^{\frac{1}{2}}$;

(4) Repeat the above procedure for a large number of subsample J, and retain the one with the lowest P_J;

(5) For this retained subsample J, compute $T(\mathbf{X}) = T_J$ and $\mathbf{C}(\mathbf{X}) = c^2(n, K)(\chi^2_{K,\,0.50})^{-1} m_J^2 \mathbf{C}_J$, where $c^2(n, K)$ is a small-sample correction term calculated as $[1+15/(n-K)]^2$ and $\chi^2_{K,0.50}$ is the median of the χ^2 distribution with K degrees of freedom.

The T($\mathbf{X}$) and $\mathbf{C}$($\mathbf{X}$) calculated in step (5) are the MVE-estimated mean and covariance matrix.

Intensive sampling and computation are required to find the solution in the MVE analysis. The total number of subsampling required depends on the values of K and n (Rousseeuw and Leroy 1987). It increases quickly with an increase in K and/or n. Based on the MVE-estimated mean T($\mathbf{X}$) and covariance $\mathbf{C}$($\mathbf{X}$), the following statistic, similar as MD^2, can be calculated,

$$W_j^2 = (\mathbf{x}_j - \mathrm{T}(\mathbf{X}))\mathbf{C}^{-1}(\mathbf{X})(\mathbf{x}_j - \mathrm{T}(\mathbf{X}))^t .$$

For a data point $\mathbf{x}_j$, if its $W_j^2 > \chi^2_{K,0.975}$, it is defined as an outlier, otherwise it is defined as a `normal' observation.

Principal component analysis

Principal component analysis is a multivariate technique for examining the relationship among several quantitative variables. Giving a data set with K numerical variables $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_K$, each of which has n individuals, PCA linearly transforms the variates $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_K$ to new variates $\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_K$. These new variates are the principal components (PC). The original data observation $\mathbf{X}^{(i)}$, which is the observation vector for the ith individual denoted as $\mathbf{X}^{(i)}=(X_{i1}, X_{i2}, \ldots, X_{iK})^t$, is transformed to corresponding PC scores $\mathbf{Y}^{(i)}=(Y_{i1}, Y_{i2}, \ldots, Y_{iK})^t$. Each PC is a linear combination of the original variables, with coefficients equal to the eigenvectors of the correlation or covariance matrix of the K variables. The principal components are sorted by descending order of the eigenvalues, which are equal to the variances of the components (Rao 1964).

Principal component analysis is often used to summarize data of multivariate nature and reduce the number of variables (Rao 1964, Cooley and Lohnes 1971). Dimensionality reduction is effective when q ($q<K$) of the components $\mathbf{Y}$ convey most of the sample information inherent in the $\mathbf{X}$. In this case the original observations $\mathbf{X}^{(i)}$ can be replaced by the first q elements of the corresponding PC scores. The number of variables measured in a fisheries or ecological study is often large. A PCA can be used to replace a large number of the original variables with a few PCs. These derived PCs are then used for further regression analyses with other variables. Such an approach of combining PCA and regression analysis is often referred to as principal component regression analysis (Hill et al. 1977, Mason and Gunst 1985, Vogt and Kolsett 1987).

A two-step procedure is proposed for developing an indicator for monitoring temporal variations in size for fish of all age groups. The robust MVE procedure is applied first to size-at-age to identify possible outliers in the data. In the next step a weighted PCA (SAS 1987) is applied to the size-at-age. For the weighted PCA, data identified as outliers in the MVE analysis are given a weight of 0 (thus effectively remove the impacts of these data on PCA) and the other "normal" data are given a weight of 1. For size-at-age data with K age groups observed for n years (or n year classes), the number of PC derived from the PCA is K. The resultant K PCs are interpreted with respect to the original size-at-age data using eigenvector values calculated from the PCA. This can establish the relationship between the PCs and original size-at-age data. If the correlation between sizes of age groups is high, the first PC, which always explains the largest proportion of variance inherent in the original data among all PCs, will be a good indicator of sizes for fish of all age groups included in the analysis. This PC can then be interpreted as an overall indicator of fish size-at-age. Temporal changes in size-at-age can be evaluated using the scores of the first PC.

An essential assumption implied in a PCA is that the relationship between variables included in the PCA is linear (Rao 1964, Manly 1991). This assumption may be violated in analyzing fish size-at-age data because the rate of growth in size tends to decrease with age and the relationship between sizes of different ages may be nonlinear, especially for long-lived fish (Ricker 1975). Two approaches can be used to avoid this problem: transforming size-at-age data (e.g. logarithm) and grouping age classes to ensure that the relationship between sizes of age classes within each group is linear. These two approaches can be used together. If grouping of age classes is used, PCA should be conducted separately for each age group.

Application

Previous studies have shown that cod in many areas of Northwest Atlantic Canada have experienced pronounced changes in growth over the last 20 years (Beacham 1983, Hutchings and Myers 1994, Lilly 1996, Shelton et al. 1996, Myers et al. 1997). Declining size-at-age was observed in some stocks during this period. However, many previous studies evaluating temporal changes in size-at-age did not consider the multivariate nature of the data. Temporal variations in cod size-at-age data were usually evaluated separately for each age group (Beacham 1983, Lilly 1996). Such an approach disregards covariance in sizes among different age groups. Because temporal patterns may differ for different age groups, inconsistency may arise when different age groups are used in evaluating temporal variation in size-at-age values.

In this study, the proposed two-step approach was applied to cod size-at-age data collected from fishery-independent bottom trawl survey in the Northwest Atlantic, NAFO subdivision 3Ps (Fig. 1). Size-at-age data were available for 20 year classes from 1971 to 1990. The among-cohort variations in size-at-age of cod were examined using the PCs derived from the proposed method. Because cod is a long-lived fish species, to avoid the problem of nonlinearity, only the first 6 age groups of size data were included in the analysis and size-at-age data were log-transformed.

Results and Discussion

Variations in size were observed among year classes included in this study for each age group (Fig. 2). Overall, the size-at-age values of the recent year classes tended to be smaller when compared with those of the 1970s year classes. However, because data were log-transformed and there were large differences in sizes among different age groups, the differences in temporal variations among age groups were difficult to be evaluated from Figure 2 which is commonly used in showing temporal variations graphically (Lilly 1996). In order to make the scales of size in different age groups comparable, log-transformed size-at-age data were standardized using the following formula:

$$y_{ij} = \frac{x_{ij} - \bar{x}_i}{S_i},$$

where x_{ij} is log-size at age i for year class j, $\bar{x}_i$ and S_i are the mean and standard deviation of logarithm sizes at age i across all year classes, and y_{ij} is standardized log-size at age i for year class j. Such a standardization did not change the patterns of temporal variation in size at age, but ensured that size-at-age data had the same scale for different age groups (Fig. 3). It is apparent that there were differences in temporal variations among age groups. Temporal variations were large for age groups 1, 2, 3 and 6, while the variations were relatively small for ages 4 and 5 (Fig. 3). In general, we conclude that the size-at-age tends to decrease for recent year classes. However, different interpretations could be derived with respect to detailed temporal variations when different age groups were used (Fig. 3).

Three year-classes, 1971, 1972, and 1974 were identified as outliers in the MVE analysis of size-at-age data. They were subsequently given a weight of 0 in PCA. Eighty-six percent of the variance in the size-at-age data was explained by the first three principal components in the PCA. The first PC explained 56% of the variance, and the second and third explained 17% and 13%, respectively. The PC4, PC5 and PC6 together only explained 14% of the variance inherent in the original size-at-age data, and were thus not important to this study.

The correlation coefficients between size-at-age variables and the first PC in the eigenvector ranged from 0.50 for age 4 to 0.30 for age 1 (Table 1). Such a small range in the correlation coefficients suggested that the first component was an overall indicator of sizes for all six age groups. The correlation coefficients between the size-at-age variables and the first PC were positive for all the six age groups (Table 1), implying that a year-class having a larger score on the first PC tended to have a larger size.

The scores on the first PC varied greatly among year classes (Fig. 4). The 1971 year-class, which was identified as an outlier in the MVE analysis, had the largest value for the scores on the first PC, indicating that the 1971 year class had the largest size prior to age 7 among the cohorts included in this study. The values of scores (thus sizes) decreased for cod from year classes 1971 to 1974, and then increased from year classes 1974 to 1979. Since the 1979 year class, the sizes decreased again. From year class 1981to the most recent year class included in this study, the sizes of cod prior to age 7 were much smaller than those for year classes in 1970s (Fig. 4).

The correlation between the second PC and size at age 2 was positive and much higher than the correlation between the second PC and other age groups (Table 1). Thus, the second PC was positively related to size at age 2. A year class with a large score on the second PC tended to have a large size at age 2. The correlation between the third PC and size at age 1 was positive and much higher than the correlation between the third PC and other age groups (Table 1). Thus, the third PC was positively related to size at age 1. A year class with a large score on the third PC tended to have a large size at age 1. The scores on both the second and third PCs fluctuated for year classes from 1971 to 1990 with no clear-cut patterns (Fig. 4). It is obvious that neither the second nor the third PC was a good indicator for sizes of all age groups, and thus cannot be used for evaluating temporal variations in size-at-age.

The MVE analysis showed that three year classes (i.e. 1971, 1972 and 1974) were outliers. This might have resulted from exceptional values in some age groups for these year classes. The 1970 year class had exceptional high values for sizes at ages 5 and 6. The 1971 year class had high values for sizes at ages 4 and 5, while the 1974 year classes had exceptionally low size at age 4. However, without knowing the level of measurement errors associated with these size data, it is difficult to tell whether these year classes were defined as outliers as a result of exceptional large measurement errors or other reasons. If it is reasonable to assume that measurement errors are more or less the same for all data included in the study, the exceptional values may result from exceptional growth or the inclusion of a large number of samples from a different stock that has different growth patterns for these three year classes (Rollet et al. 1995, Lilly 1996).

Different hypotheses have been developed to explain the decrease in size-at-age observed in cod populations in recent years. These hypotheses include large scale temporal variations in water temperature (Beacham 1983, Hutchings and Myers 1994, Gomes et al. 1995), changes in stock biomass (Hanson and Chouinard 1992, Swain 1993), stock overfishing (Trippel 1998) and variation in prey species biomass (Krohn et al. 1997). Regardless of the factors causing the decrease in size-at-age, most authors suggest that such a phenomenon is indicative of population stress (Kovstsova 1995, Trippel 1995, 1998). If observations/data are available for these environmental variables, the overall indicator for size-at-age identified in the robust PCA can be used as a variable representing fish size in regression analysis with these environmental variables. Such a principal component regression analysis can identify whether the temporal variations in fish size are related to the temporal variations of environmental variables (Hill et al. 1977, Vogt and Kolsett 1987). This approach can reduce the number of variables without the loss of information inherent in the original variables. When size-at-age data are directly used in an analysis with environmental variables, an arbitrary decision has to be made in determining which age group should be included in the analysis. Information of other age groups that are not included in the analysis is thus lost.

The size-at-age-1 data were obtained from low sample sizes (Lilly 1996), and were thus less reliable. However, whereas this may undermine any reliable interpretation using such data series in a conventional univariate analysis, it is less disruptive to the method proposed in this study because it assesses the cumulative effect of size-at-age over the total period of time the cohorts are considered in the study. The ability to minimize disruptions from data quality or availability issues is an important strength of the methodology proposed in this study.

Acknowledgement

We would like to thank Dr. George Lilly and Dr. Peter Shelton of the Northwest Atlantic Fisheries Research Centre for allowing us to use their data and providing insights about the cod fishery in 3Ps. Financial support of this study was provided by an NSERC-Industry Chair Program in Fisheries Conservation and an NSERC operating grant to Y. Chen and a Department of Fisheries and Oceans National Scholarship in Ocean Studies to L. Mello.

References

Beacham, T. D. 1983. Growth and maturity of Atlantic cod (*Gadus morhua*) in the southern Gulf of St. Lawrence. Can. Tech. Rep. Fish. Aquat. Sci. 1142.

Beverton, R. J. H., and S. J. Holt 1957. On the dynamics of exploited fish populations. Fish. Invest. Ser. 2 Mar. Fish. G.B. Minist. Agric. Fish. Food No. 19.

Bowering, W.R., W. B. Brodie, and M. J. Morgan 1996. Changes in abundance and certain population parameters of American plaice on St. Pierre Bank off Newfoundland during 1972-1994, with implications for fisheries management. N. Am. J. Fish. Manange. 16: 747-769.

Campbell, N. A. 1980. Robust procedures in multivariate analysis I: robust covariance estimation. Applied Statistics29:231-237.

Charnov, E. 1993. Life history invariants. Oxford University Press, New York.

Chen, Y. and H. H. Harvey 1994. Maturation of white sucker, *Catostomus commersoni*, populations in Ontario. Can. J. Fish. Aquat. Sci. 51:2066-2076.

Chen, Y. and H. H. Harvey 1995. Growth, abundance, and food supply of white sucker. Trans Am. Fish. Soc. 124:262-271.

Chen, Y., D. A. Jackson, and J. E. Paloheimo 1994. Robust regression approach to analyzing fisheries data. Can. J. Fish. Aquat. Sci. 51:1420-1429.

Chen, Y. and J. E. Paloheimo. 1998. Can a more realistic model error structure improve the parameter stimation in modelling the dynamics of fish populations? Fish. Res. 38: 9-17.

Cooley, W. W. and P. R. Lohnes. 1971. Multivariate Data Analysis. John Wiley &Sons, New York.

Devlin, S. J., R. Gnanadesikan, and J. R. Kettenring 1981. Robust estimation of Dispersion matrices and principal components. J. Am. Stat. Ass. 76:354-362.

Gomes, M. C., R. L. Haedrich and M. G. Villagarcia. 1995. Spatial and temporal changes in the groundfish assemblages on the northeast Newfoundland/Labrador Shelf, northwest Atlantic, 1978-1991. Fish. Oceanogr. 4: 2, 85-101.

Hanson, J. M., and G. A. Chouinard 1992. Evidence that size-selective mortality affects growth of Atlantic cod (*Gadus morhua* L.) in the southern Gulf of St. Lawrence. J. Fish. Biol. 41:31-41.

Hilborn, R., and C. J. Walters 1992. Quantitative Fisheries Stock Assessment: Choice, Dynamics & Uncertainty. Chapman & Hall, N.Y.

Hill, R. C., T. H. Fomhy, and H. H. Johnson. 1977. Component selecting norma for principal component regression. Common Stat. Theor. Math. 6:300-331.

Hutchings, J. A., and R. A. Myers 1994. Timing of cod reproduction: interannual variability and the influence of temperature. Mar. Ecol. Prog. Ser. 108:21-31.

Jackson, D. A. 1993. Stopping rules in principal component analysis: a comparison of heuristical and statistical approaches. Ecology 74:2204-2214.

Jensen, A. L. 1996. Beverton and Holt life history invariants result from optimal trade-off of reproduction and survival. Can. J. Fish. Aquat. Sci. 53:820-822.

Kovtsova, M. V. 1995. Changes in growth and maturation of the Barents Sea plaice (*Pleuronectes platessa* L.) in 70s-90s. ICES Council Meeting Papers 12. ICES,Copenhagen (Denmark), 12 pp.

Krohn, M., S. Reidy, and S. Kerr 1997 Bioenergetic analysis of the effects of temperature and prey availability on growth and condition of northern cod (*Gadus morhua*). Can. J. Fish. Aquat. Sci.54 (Suppl. 1):113-121.

Krzanowski, W. J. 1988. Principal of Multivariate Analysis. Clarendon Press, Oxford, UK

Lilly, G. R. 1996. Growth and condition of cod in Subdivision 3Ps as determined from trawl surveys (1972-1996) and sentinel surveys (1995). DFO Atlantic Fish. Res. Doc. 96/69.

Manly, B. F. J. 1991. Randomization and Monte Carlo Methods in Biology. Chapman and Hall, New York.

Mason, R. L. and R. F. Gunst. 1985. Selecting principal components in regression. Stat. Prob. Lett. 3:299-301.
Moreau, J. 1987. Age and growth of fish, p. 101-143. In R.C. Summerfelt and G.E. Hall [ed] Fish Growth. Iowa State University Press, Ames, Iowa.
Myers, R. A., G. Mertz, and P. S. Fowlow 1997. Maximum population growth rates and recovery times for Atlantic cod, *Gadus morhua.*
Nikolskii, G. V. 1965. Theory of fish population dynamics. Oliver & Boyd, Edinburgh, U.K.
Paloheimo, J. E. and L. M. Dickie. 1965. Food and growth of fishes I. A growth curve derived from experimental data. J. Fish. Res. Bd. Can. 22:521-542.
Rao, C. R. 1964. The use and interpretation of pricipal component analysis in applied research. Sankhya A, 26:329-358.
Ricker, W. E. 1975. Computation and Interpretation of Biological Statistics of Fish Populations. Fish. Res. Board Can. No. 191
Roff, D. A. 1984. The evolution of life history parameters in telosts. Can. J. Fish. Aquat. Sci. 41:989-1000.
Rollet, C., J-C. Brêthes and A. Frechet 1995. Spatial and temporal variation in cod length frequencies and length at 50% maturity in Divisions 3Pn, 4RS and 3Ps. DFO Atl. Fish. Res. Doc. 95.103.
Rousseeuw, P. J. 1984. Least median of squares regression. J. Am. Stat. Ass. 79:871-880.
Rousseeuw, P. J. 1985. Multivariate estimation with high breakdown point. In Mathematical Statistics and Applications (Vol. B), Eds. W. Grossmann, G. Pflug, I. Vincze, and W. Wertz, Dordrecht: Reidel Publishing 283-297.
Rousseeuw, P. J., and A. M. Leroy 1987. Robust regression and outlier detection. John Wiley & Sons Inc. New York, N. Y.
Rousseeuw, P. J. and B. C. van Zomeren. 1990. Unmasking multivariate outliers and leverage points. J. Am. Stat. Ass. 85:633-644.
SAS. 1987. SAS stat. Guide for personal computers. SAS Institute, Cary, NC.
Shelton, P. A., D. E. Stansbury, E. F. Murphy, J. Brattey, and G. R. Lilly 1996. An assessment of the cod stock in NAFO subdivision 3Ps. DFO Atl. Fish. Res. Doc.96/91.
Swain, D. P. 1993. Age- and density-dependent bathymetric pattern of Atlantic cod (*Gadus morhua*) in the southern Gulf of St. Lawrence. Can. J. Fish. Aquat. Sci. 50:1255-1264.
Trippel, E. A. 1995. Age at maturity as stress indicator in fisheries. SioScience 45:759-771.
Trippel, E. A. 1998. Egg size and variability and seasonal offspring production of young Atlantic cod. Trans. Am. Fish. Soc. 127:339-359.
Vogt, N. B. and K. Kolsett. 1987. Composition activity relationships _CARE. Part III. Polynomial principal component regression and response surface analysis of mutagenicity in air samples, 1981. Report 87:747, Chemometrics and Indelligent Lab Systems.
Walters, C. J. 1999. Evaluation of quota management policies for developing fisheries. Can. J. Fish. Aquat. Sci. (in press).

Table 1. Eigenvectors for the first three components in the principal component analysis of logarithmic sizes at ages 1 to 6 for cod in 3Ps.

	Principal component		
Log size at age	PC I	PC II	PC III
1	0.30	0.03	0.90
2	0.45	0.41	0.07
3	0.32	0.70	-0.31
4	0.50	-0.16	-0.11
5	0.45	-0.34	-0.11
6	0.40	-0.46	-0.25

Figure 1. Map of the Northwest Atlantic Fisheries Organization (NAFO) fishing areas in the Newfoundland and Labrador shelf.

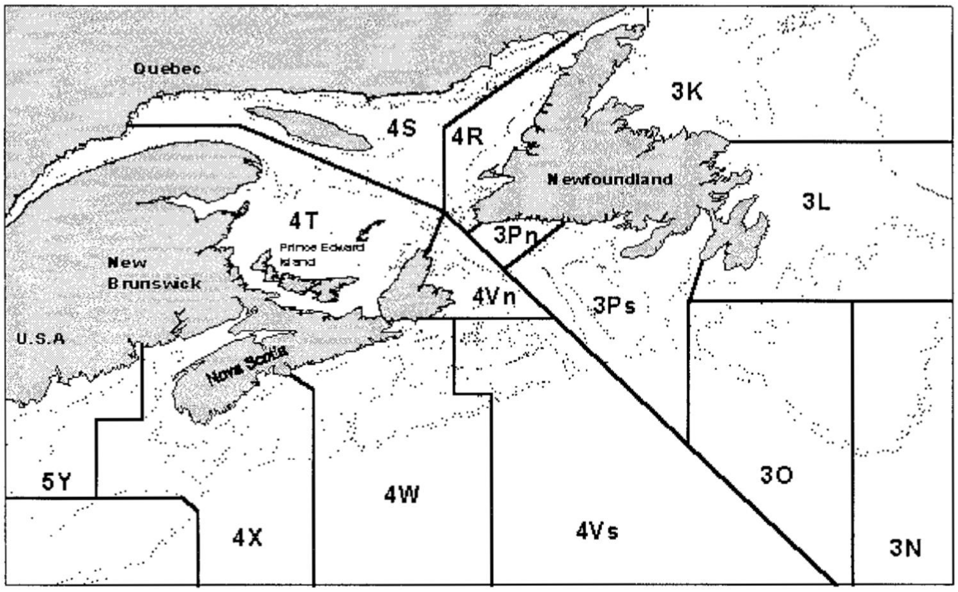

Figure 2. Variations in size of six age groups for the 1971 to 1990 year classes.

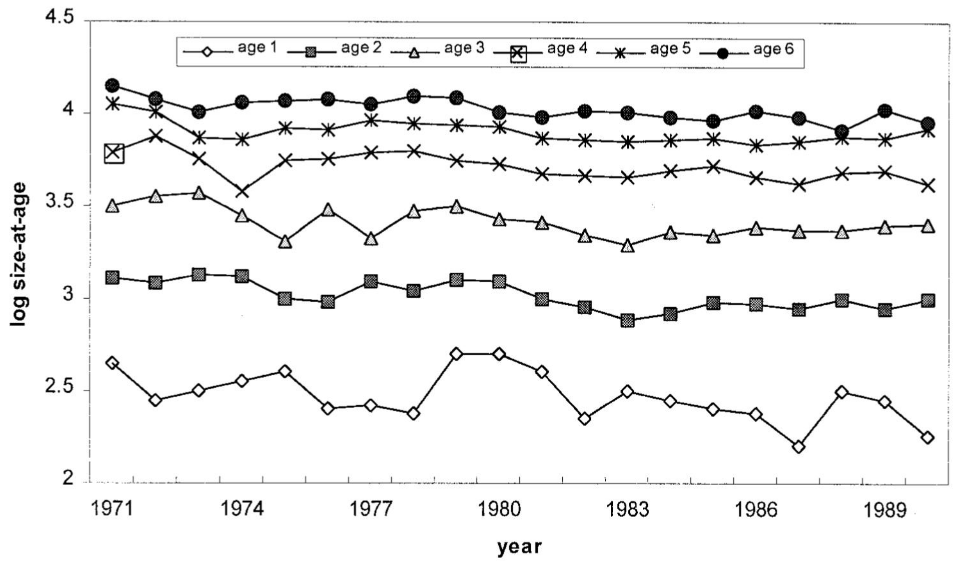

Figure 3. Plot of standardized size-at-age data for six age groups of 20 year classes from 1971 to 1990.

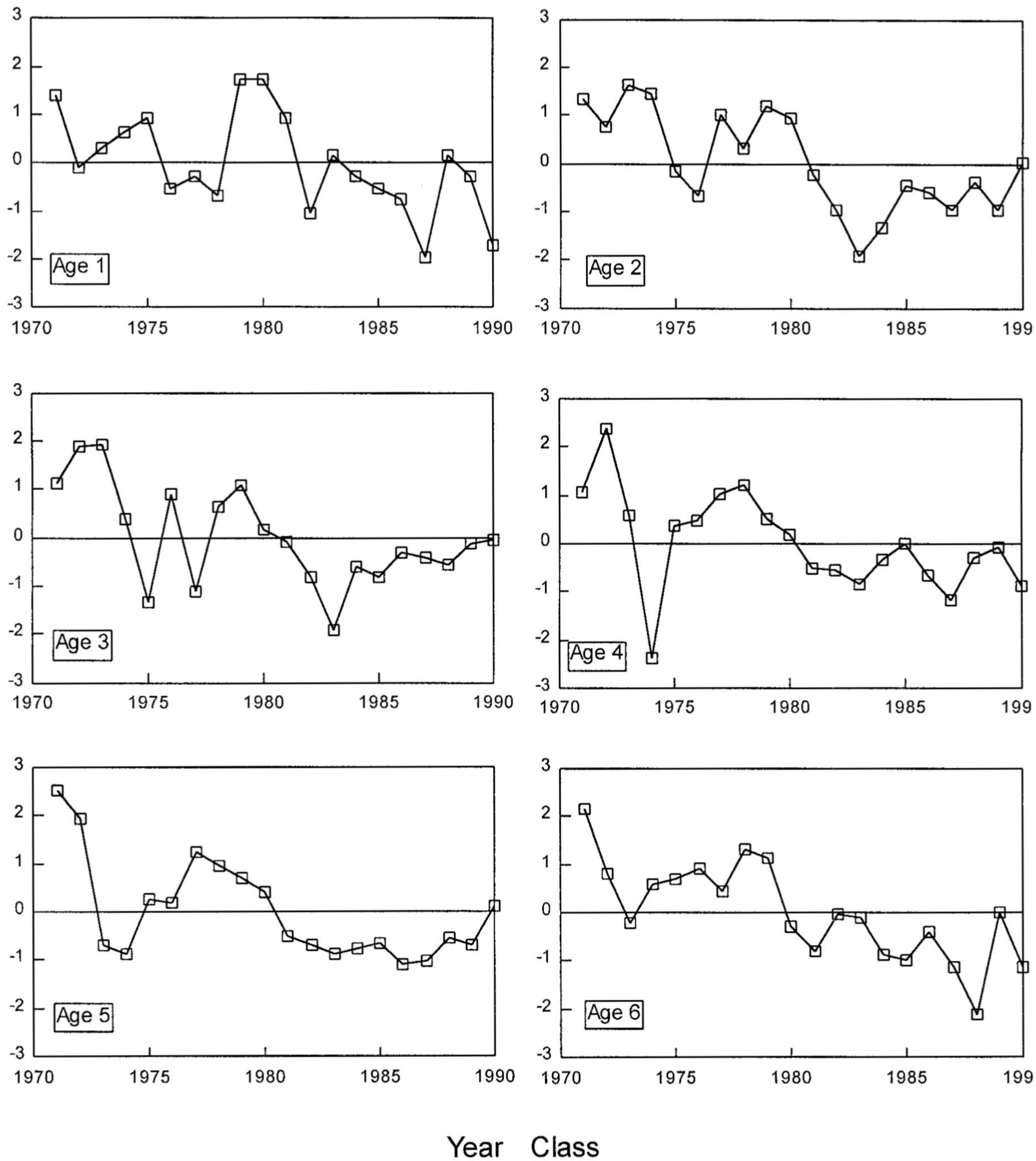

Figure 4. Plot of the first three principal components (PC) derived in the robust principal component analysis of size data of age groups 1 to 6 for the 1971 to 1990 year classes.