International Council for the Exploration of the Sea

C.M.1989/K:5
Ref.: D

REPORT OF THE WORKSHOP ON THE MULTIVARIATE ANALYSIS OF SHELLFISH STOCKS

Lowestoft, UK, 4-7 July 1989

*General Secretary
ICES
Palægade 2-4
DK-1261 Copenhagen K
DENMARK

T A B L E O F C O N T E N T S

---ooOoo---

1. INTRODUCTION

1.1 Participation

Members

C. Bannister (Chairman)	UK (England)
N. Bailey	UK (Scotland)
D. Bennett	UK (England)
R. Briggs	UK (N. Ireland)
G. Conan	Canada
M. Fogarty	USA
P. Hillis	Ireland
C. Hopkins	Norway
Y. Morizur	France
S. Munch-Petersen	Denmark
S. Murawski	USA
M. Nicholson	UK (England)
E. Nilssen	Norway
Y. Simard	Canada
J. Volstad	Norway

Observers

J. Addison	UK (England)	D. Hogarth	UK (England)
F. Almeida	USA	J. Idoine	USA
S. Chang	USA	A. Lange	USA
Y. Chiasson	Canada	W. Michaels	USA
S. Clark	USA	M. Moriyasu	Canada
M. Comeau	Canada	B. O'Gorman	USA
R. Conser	USA	T. Polacheck	USA
C. Darby	UK (England)	C. Rocha	UK (England)
S. Edwards	USA	A. Thompson	UK (England)
W. Emerson	USA	E. Wade	USA
J. Forrester	USA	G. Waring	USA
K. Foster	USA	P. West	UK (England)
R. Fryer	UK (Scotland)		

1.2 Terms of Reference

Council Resolution 1988/2:18 states: "A Workshop on Multivariate Analysis of Shellfish Stocks will be held under the Chairmanship of Dr C. Bannister (UK) in Lowestoft from 4-7 July 1989 at national expense to study the general application of multivariate analysis to population data for *Nephrops*, *Cancer*, *Homarus*, *Chionocetes*, and *Pandalus*.

1.3 Meeting Arrangements

Formal proceedings were convened by the Chairman at the Fisheries Laboratory, Lowestoft, UK, from 4-7 July 1989. These were preceded by an informal preparatory meeting convened with the assistance of Dr S. Murawski at the North-East Fisheries Center, Woods Hole, Mass. USA, from 8-11 May. These arrangements facilitated a wide participation of USA and Canadian members and observers who were unable to travel to Lowestoft. Linkage and continuity were achieved by the attendance of Dr Bannister and Mr Nicholson at both meetings.

1.4 Origin, Aims and Structure of the Workshop

The Workshop was proposed by the Shellfish Committee at the 76th Statutory Meeting, Bergen, 1988, in response to the presentation of the Report of the Study Group on *Nephrops* (ICES, Doc.C.M.1988/K:29), Coleraine, Northern Ireland.

This report discussed the time-space difference in *Nephrops* size distribution data both within and between different fisheries. It raised the question whether the use of multivariate techniques might help to clarify the relative effects on size distribution and abundance of such factors as rate of exploitation, stock density, sediment type, depth and diel behaviour. The Shellfish Committee decided to promote a multivariate approach to the *Nephrops* work, but also to consider applications to the other species contained in the terms of reference, in order to promote a wider awareness of the potential benefits of using this type of analysis. In the event, the Chairman of the Workshop asked members for as wide a range of examples of application as possible, and deemed the practical aims of the Workshop as being:

1) To present, discuss or undertake the application of multivariate analysis to a wide range of shellfish examples and problems.

2) To foster an interchange of experience of multivariate analysis between statistical advisors, established biological users and beginners unfamiliar with the methods.

3) To promote good statistical thinking and biological interpretation in the use of multivariate techniques with shellfish data.

4) To exchange views and information about user experience of different statistical packages.

The activities of the Workshop fell into the following broad categories:

- Overview of the objectives and scope of multivariate analysis.

- Presentation of an ordered series of case histories to illustrate and permit discussions of applications.

- Live analysis of representative data sets brought to the meeting.

- Discussion of analytical guidelines.

- Exchange of views on software applications.

- Discussion of contingent topics such as sampling, transformation, and the analysis of response surfaces.

- Ecological applications.

2 THE SCOPE OF MULTIVARIATE ANALYSIS

Overviews were given by Janice Forrester (Woods Hole) and Mike Nicholson (Lowestoft).

Very generally, the multivariate problem is one in which the fundamental biological relationships may be obscured by the difficulty of representing and interpreting data in more than three dimensions. The analytical task is to reduce these data by some manipulation which identifies patterns or associations likely to have biological or distributional meaning, or to represent character sets which permit rigorous discrimination between data from different sources with known precision. Multivariate methods therefore pursue data reduction, and sorting, grouping or classification. The main approaches of interest in this Workshop were as follows:

1) The systematic association of multivariate observations to groups whose members are closer to each other than to other observations or groups (cluster analysis, CLA).

2) The stepwise determination of the set of linear combinations of the variables which allow as much as possible of the multidimensional variation to be represented in two dimensions (principal component analysis, PCA).

3) Calculation of linear combinations of the variables which, projecting the data for known groups onto new axes, achieves the largest ratio of the between group variation to the within group variation, thus leading to an objective criterion for subsequent classification (discriminant function analysis, DFA).

4) An extension to PCA is the technique of canonical correlation, which finds for each of 2 sets of variables that linear combination of each set which maximises the correlation between them.

The fundamental framework of statistical analysis is the construction and testing of *a priori* hypothesis. In this Workshop, however, the main approach was primarily that of data exploration, looking for underlying biological and environmental relationships in existing data sets. The results might be tested in subsequent more controlled experiments. Exceptions to this exploratory approach were the examples using morphometric data to test specific hypotheses about functional and morphometric maturity, and the analyses of *Arctica* and *Pandalus* based on existing publications. (see later).

The emphasis throughout was on the presentation and description of the data; the underlying rationale for the analysis and the statistical interpretation; and on the biological interpretation.

The statistical methods were generally discussed in a heuristic way, avoiding becoming trapped in unnecessary mathematical detail. Attention was given however to the interpretation of the correlation matrix, the role of the covariance matrix, options for transforming and recentering the data, and the identification of the new variables contributing to the bivariate presentation of the data following rotation.

3 CASE STUDIES AND EXAMPLES

The case studies and examples were assembled, presented and discussed on the basis of three criteria, viz., species, problem type and analytical method. Table 1 lists these examples by species, problem and author, and Table 2 by problem and analytical method. In Table 1 asterisked items are those examples where data were re-analysed at the Workshop. Remaining items were represented in detail, but were based on work previously published, or previously carried out, but either published informally or not yet published.

For most of these examples the Workshop was able to assemble a control sheet and a dossier. The control sheet described the problem, the available data, technical aspects of the analysis, the results, and the biological interpretation. The dossier included, where available, or desirable, output tables of the correlation and/or covariance matrix, the principal component or discriminant function scores, and illustrations showing the original variables and the bivariate projections of new variables.

An example of the control sheet is included in Appendix 1, illustrating the approach for gulf of st lawrence morphometric study of Conan and Comeau (1986). The volume of material in the dossiers is too large to included in this report. Taken together, however, they represent an informative survey of the major steps and results involved in applying multivariate methods to the shellfish problems dealt with. A brief summary of the main results is included here as Appendix 2, but the value of the Workshop would be enhanced if the dossiers could be produced in full in the form of, say, a Cooperative Research Report.

4 GUIDELINES

Various operational questions and analytical steps emerged during the course of the Workshop and it is tempting to see these prescriptively, and to forget that the art of statistical analysis is as much to adapt the analysis to the problem as to adapt the problem to a specific analytical technique. Even so, some general features common to all analyses are a preliminary screening of the data; a confirmation or otherwise of preliminary expectation about the data; a discovery of new and possibly unexpected features; a basic summary in statistical format; and the possible development of more formal analyses. Within this broad framework we tend to ask some "guideline" questions and these are summarised below in case they may be of general interest.

Initial Questions

What is the biological problem (in the sense of Table 2)?

What variables are available and what biological relationship do we expect them to have?

What multivariate method seems appropriate? (see below)

What specific variables will be used, and do they need prior transformation?

Problem - Method Pathway

1) Are we looking for structure in the data, without much prior knowledge of what to expect?

   Consider preliminary analysis by cluster analysis, which may assist in identifying groupings based on statistical distance criteria.

   Could the data be filtered using a dissimilarity index?

   Which linkage algorithm is most appropriate?

   What dissimilarity criterion is to be adopted in making classification?

   Are the results robust for different linkage criteria?

   Can the results of this subjective analysis be tested independently?

2) Is there an expected data structure which can be used for classification?

   Apply discriminant function analysis to test for the variables most useful in achieving discrimination.

   What *a priori* classification is to be used for the test data?

   What level of precision is acceptable/necessary in making the classification (type I/type II errors).

   Are all the variables needed to achieve this discrimination (check the discrimininant scores).

3) Is it necessary to achieve some reduction in the number of dimensions in order to investigate and identify relationships?

   Carry out principal component analysis.

   Examine the correlation matrix and the individual bivariate plots. (If data are not initially correlated, there is no point in continuing).

   Is the analysis to be done with the correlatrion or the covariance matrix? Always evaluate the covariance structure.

Does the analysis achieve reduction, i.e., can most of the variation be represented by the new orthogonal variables in one or two bivariate projections?

Is it possible to identify outliers, and show what variables are correlated with the principal axes? (Examine the principal component scores).

Are any of the variables redundant?

Are the projections of the new variables homogenous, or is there evidence for interpretable heterogeneity or grouping?

Note that as the results of PCA are scale dependent, different results will be given by standardised variables.

General Points

Computer packages make a cookery book approach technically feasible, but interpretation of results usually requires biological knowledge based on initial analysis of the data and the systems under study.

PCA and DFA identify associations and correlations but do not establish cause and effect.

Populations are subject to time-dependent change and "single survey" analyses may be misleading.

Simplification following data reduction should allow the planning of rigorous experiments or more simplified data collection to test hypotheses.

Multivariate analysis provides a useful way of validating sampling stratification.

Try more than one method - are results robust for different methods and data treatments?

The most useful role of these techniques is for exploring and screening data so have fun with this approach!

Trawl Surveys

When analysis population characteristics derived from trawl surveys ordination techniques assume independence of the individual observations. As noted by Digby and Kempton (1987), "with direct methods of ordination, the experimenter must specify the environmental factors of interest, and have independent knowledge of the score at each site for each factor (for species ordination) or the species response to each factor (for site ordination)". For trawl or dredge surveys the independent unit of observation is the trawl haul, not the individuals of the species measured from the haul. If the population aggregates by sex, size or age, there may be an intra-haul correlation for these characteristics, in which case individuals samples within the tow will not necessarily be independent observations. The degrees of freedom are related to the number of tows rather than the individuals in the sample, and it is preferable to have many short tows rather than a few long tows. Note also that the comparison of length distributions by such as the Kolmogorov-Smirnov test also assumes independent measurement of individuals.

5 STATISTICAL PACKAGES

User experience and views were exchanged on the following packages relevant to the practice of multivariate analysis. Methods in brackets are those discussed at the meeting in relation to each package.

Minitab VI and VII	(PCA, DFA)
Genstat 5	(CLA, PCA, DFA)
SAS	(Full range of techniques)
NT Syst	(PCA. Multidimensional scaling)
Addad	(CLA, PCA, DFA. Correspondence and factor analysis)
Clustan	(CLA)
Systat II	(CLA; multiple general linear hypothesis testing; multidimensional scaling).
Twinspan-Decorama	

6 BIBLIOGRAPHY

A bibliography of selected multivariate papers in aquatic science was compiled by E. Nilssen (Norway) and circulated at the meeting.

7 CONCLUSIONS

The Workshop stimulated the preparation and presentation of 16 different case studies illustrating the role of multivariate statistical techniques in several important types of biological analysis. These studies included new analyses carried out at the Workshop on *Pandalus borealis* data for the Gulf of Maine and the North Sea; *Nephrops norvegicus* data for the Irish Sea and West of Scotland; and *Spisula solidissima* data for eastern USA. There was an effective and user-friendly exchange of views concerning multivariate analysis which succeeded in demonstrating the utility of these methods to non-practitioners. This is a very practical way of upgrading scientific standards in a problem-orientated setting and it is recommended that the Shellfish Committee pursue the Workshop approach with other biological, assessment and statistical topics when relevant or practicable. The Shellfish Committee is asked to endorse the Workshop request that the full report be published in the Cooperative Research Report series.

8. REFERENCES

In addition to case studies drawn from published work, which is referenced here, members contributing at the Workshop to new but as yet unpublished case studies are referenced as below, albeit in unconventional form.

Bailey, N. and R. Fryer (unpublished pers. comm.). A preliminary multivariate analysis of trawl survey and environmental data for Nephrops norvegicus from three areas in the west of Scotland.

Clarke, S., J. Forrester, and M. Nicholson (unpublished pers. comm.). A preliminary multivariate analysis of abundance and environmental data collected during trawl surveys for Pandalus borealis in the Gulf of Maine, USA.

Conan, G.Y. and M. Comeau. 1986. Functional maturity and terminal moult of male snow crab Chionocetes opilio. Can. J. Fish. Aquat. Sci. 43: 1710-1719, 32 pp.

Conan, G.Y., M. Comeau, and M. Moriyasu. 1985. Functional maturity of the American lobster Homarus americanus. ICES, Doc. C.M.1985/K:29.

Digby, P.G.N. and B.A. Kempton. 1987. Multivariate Analysis of Ecological Communities. Chapman and Hall, London.

Fogarty, M.J. 1981. Distribution and relative abundance of the ocean quahog Arctica islandica in Rhode Island Sound and off Martha's Vineyard, Mass. Journal of Shellfish Research, 1, no. 33-39.

Hillis, J.P. (unpublished pers. comm.). A preliminary multivariate analysis of abundance, length and environmental data for Nephrops norvegicus collected during a 1988 research vessel survey in the Irish Sea.

Hogarth, D. (unpublished pers. comm.). Observations on the problem of identifying the size of maturity in a population of Paralomis granulosa from the Falkland Islands.

Hopkins, C.C.E., E.M. Nilssen, A. Hermannsen, B. Vajaa, and L. Dabso (unpublished pers. comm.). Body composition and energetic predictors of age/season in the prawn Pandalus borealis.

Latrouite, D. and Y. Morizur. 1988. Analyse quantitative et qualitative des captures de tourteau Cancer pagurus par la flottille de Manche-Ouest. ICES, Doc. C.M.1988/K:33.

Munch-Petersen, S., E.M. Nilssen, and C.C.E. Hopkins (unpublished pers. comm.). Pandalus borealis in the North Sea. A multivariate exercise in the separation of "stocks".

Murawski, S. (unpublished pers. comm.). A trial examination of length frequency data for Spisula solidissima in the Gulf of Maine using cluster analysis.

Murawski, S. et al. (unpublished pers. comm.). A trial investigation of differences between Spisula solidissima populations in new Jersey and Maryland, USA, using a multivariate analysis of morphological data.

Nilssen, E.M. and C.C.E. Hopkins (unpublished pers. comm.). A community analysis of the Barents Sea using multivariate methods.

Nilssen, E.M. and C.C.E. Hopkins. 1986. Biological variation between and within populations of the deep water prawn Pandalus borealis in two north Norwegian fjords. ICES, Doc. C.M.1986/K:15.

Rocha, C. (unpublished pers. comm.). Observations on the determination of the size of maturity in female Panulirus laevicanda in Brazil.

Simard, Y. and L. Savard (unpublished manuscript). Use of multivariate statistical methods to study the spatial organisation of length frequency data of Pandalus borealis in the Gulf of St. Lawrence.

APPENDIX 1

Example of Control Sheet for Snow Crab Morphometric Analysis

Title

Morphometric maturity of male snow crab, *Chionocetes opilio*, in the Gulf of St Lawrence, Canada. (Conan and Comeau, 1986).

Biological Problem

To identify the size of morphometric maturity using morphometric data.

The Data

Carapace length and width; chela length, height and width; length of second pereiopod mereopodite. Measurements made for representative sample across the size range captured by *Nephrops* trawl and commercial traps

Analysis

i) Preliminary bivariate plots using both individual characters and a composite index (Powels index).

ii) Preliminary PCA on 6 variables using data log transformed to linearise the allometric relation and stabilise the variance. Data were centred and reduced.

iii) Rejection of redundant variables and repeat of PCA on 4 variables.

iv) Use of independent observations projected onto the bivariate plane of the first and second principal components to identify variables.

v) Bivariate discriminant analysis on 2 variables (chela height, carapace width

Results

i) Log log plots of chela height or the Powles index, against carapace length indicate two overlapping sets of points potentially identifiable with a maturity difference.

ii) In a PCA with 6 variables the relevant axes are axis 1 (98% of the variance) and axis 2 (1.14% of the variance).

iii) The PCA scores of axis 1 show redundancy, i.e., carapace can be represented by either length or height, and chela by either height or length.

iv) A PCA for 4 variables shows two fields of points on axis 1 and 2 which can be identified as adults or pre-mature adults by projecting independent observations.

v) Using carapace width and chela height a discriminant function analysis on data assigned to immature and mature classes achieves nearly 100% discrimination between discriminant scores, the respective covariance matrices being significantly different at $p = 0.01$.

Biological Interpretation

Male snow crab show a clear difference in the size of the male chela at morphometric maturity. This is associated with the behavioural act of the male gripping the female at maturity. The result, which allows complete discrimination, invokes a substantial discussion about the snow crab life cycle and the differences between functional, morphometric and physiological maturity in the males of this species. There may also be some indication of time dependent differences in data obtained as exploitation of the stock advances.

Table 1 List of Workshop Examples arranged by species, problem type and author. (*Analyses carried out at the Workshop.)

Species	Problem	Source
Arctica islandica (surf clam)	Distribution in relation to sediment off Massachusetts	Fogarty, 1981
Cancer pagurus (Brown crab, tourteau)	Spatial heterogeneity and sampling stratification in English Channel	Latrouite and Morizur, 1988
Chionocetes opilio (Snow crab)	Maturity using morphometric data for Gulf of St Lawrence	Conan and Comeau, 1986
Homarus americanus (American lobster)	Maturity using morphometri data for Gulf of St Lawrence	Conan, Comeau and Moriyasu, 1986
Nephrops norvegicus (Prawn)	Time space heterogeneity in relation to sediment and depth in Irish Sea	*Hillis (pers. comm.)
	Time space heterogeneity in relation to sediment and depth in West Scotland	*Bailey and Fryer (pers. comm.)
Pandalus borealis (Deep water prawn)	Stock identification in the North Sea	*Munch-Petersen, Nilssen and Hopkins (pers. comm.)
	Spatial organisation in the Gulf of St Lawrence	Simard and Savard (pers. comm.)
	Time-space heterogeneity in Barents Sea trawl survey data	Nilssen and Hopkins (pers. comm.)
	Time-space heterogeneity at sampling stations in Norwegian fjords	Nilssen and Hopkins, 1986
	Heterogeneity in relation to temperature and depth in the Gulf of Maine	*Clarke, Forrester *et al*. (pers. comm.)
	Age/season aspects of growth and condition in the Barent Sea	Hopkins, Nilssen *et al*. (pers. comm.)
Panulirus laevicanda (Spiny lobster)	Maturity using morphometric data in Brazil	Rocha (pers. comm.)
Paralomis granulosa (Blue crab)	Maturity using morphometric data in the Falkland Islands	Hogarth (pers. comm)
Spisula solidissima	Stock identification from shell morphometrics in eastern USA	*Murawski (pers. comm.)
	Spatial distribution in eastern USA	*Murawski (pers. comm.)

Table 2 List of Workshop Examples Arranged by Problem Type and method of Analysis

Problem	Example Method	Analytical (see footnote)
Morphometric Maturity	*Chionocetes opilio* in Gulf of St Lawrence	PCA, DFA
	Homarus americanus in Gulf of St Lawrence	PCA
	Panulirus laevicanda in Brazil	PCA
	Paralomis granulosa in Falkland Islands	-
Stock Identification	Shell morphometrics of *Spisula solidissima* off New Jersey and Maryland, USA	PCA, DFA
	Length frequency data for North Sea *Pandalus borealis*	PCA, DFA
Spatial Aggregation	*Pandalus borealis* length frequency data for Gulf of St Lawrence	CPA, SV
	Spisula solidissima length frequency data for eastern USA	CPA
Time-Space Heterogeneity	Barents Sea trawl survey data for *Pandalus borealis* and six fish species	PCA
	Pandalus borealis length frequency data in some Norwegian fjords	PCA
	Sub-area differences in length frequency and sex ratio for *Cancer pagurus* in English Channel	PCA, CrA
	Abundance in relation to sediment in *Arctica islandica*	PCA, DFA
	Length frequency distribution in relation to depth and sediment in Irish Sea *Nephrops*	CrA
	Length frequency distribution and abundance in relation to depth, sediment and other factors in West of Scotland *Nephrops*	PCA
	Abundance and size in relation to depth and temperature for Gulf of Maine *Pandalus borealis*	PCA
Body Composition and Growth	Age/season aspects of growth and condition in *Pandalus borealis* of Barents Sea	PCA, MRA

Abbreviations:
CaC - Canonical Correlation
ClA - Cluster Ananalysis
CrA Correspondence Analysis
DFA - Discriminant Function Analysis
MRA - Multiple Regression Analysis
PCA - Principal Component Analysis
SV - Spatial Analysis by Variogram

APPENDIX 2

Summary Results of Case Studies and Examples

A. MORPHOMETRIC MATURITY

The aim is to identify maturity groups based on external morphological differences. These provide quick non-destructive identifiers for use in the field by fishermen, so that maturity group can, if necessary, be protected by specific management measures.

1. Gulf of St Lawrence snow crab, Chionocetes opilio

   The relation between claw dimension and body size was analysed by PCA using 6, then, because of redundancy, 4 variables. Data were log transformed, centred and reduced. There was a recognisable separation into immature and mature males in projections of the new variables onto the first and second principal axes, with the difference, separation being along axis 2. A corresponding discriminant function analysis achieved extremely good separation. The morphometric character in morphometrically mature males, a large claw, is a strong feature developed to allow a male to grip the female during copulation. Conclusions from this analysis featured an extensive discussion of the difference between physiological, morphometric and functional maturity.

2. Gulf of St Lawrence Homarus americanus

   The relation between claw dimension and body size was analysed by PCA for six variables using log transformed data centred and reduced. Male and female data were discriminated, the best discriminants being abdomen width for female, and crusher claw thickness for male. Within the sexes there was no discrimination of groups which could be used to identify maturity. In this species any morphometric changes associated with maturity must occur gradually from the juvenile stage onwards.

3. Panulirus laevicanda in Brazil

   Maturity of female spiny lobster was analysed by PCA using data on body dimension and body length. Virtually complete discrimination between immature and mature females was achieved using abdomen width.

4. Paralomis granulosa in Falkland Islands

   As with Homarus americanus samples of claw size and body size so far show no distinctive sub-groupings useful for identifying the onset of maturity.

B. STOCK IDENTIFICATION

Sedentary shellfish species often appear to have distributions which, albeit locally patchy, are essentially continuous. It can then be difficult to see an obvious basis for describing stock differences so that the management unit is then often defined as the area traditionally occupied by a "fishery" based on one or more ports. A more rigourous approach would be to look for a biological pattern by applying multivariate techniques to survey data on abundance, size distribution and morphometrics, in the hope of identifying clusters or groups which can be studied in more detail.

1. Spisula solidissima in the eastern USA

A subset of morphological data comprising shell length, width, height, weight and age were compared for two areas, New Jersey and Maryland, by PCA, to test the hypothesis of morphological differences. Results indicate a lack of substantive difference in the attributes between areas except for shell width, which may be correlated with clam density. Further analysis of this aspect of the data seems warranted.

2. Pandalus borealis in the North Sea

Single centimetre group length frequency data from three different North Sea areas were log transformed, treated as 42 variables, and analysed by PCA. In one of the source areas, the Norwegian Deeps-Skagerrak, the distribution of Pandalus is essentially continuous. Projections of the new variables on the first two principal axes identified groupings differing in size and sex. Data from the Skagerrak-Norwegian Deeps overlapped but were separate from the Fladen and Farne areas. A discriminant function analysis gave a text book separation between Fladen, Norwegian coast and Skagerrak data due to biological differences strongly suggestive of stock differences. This preliminary analysis is worth further detailed investigation by specific data collection.

C. SPATIAL AGGREGATION

Within a geographic area shellfish populations may aggregate in local patches because of particular requirements at settlement, such as depth and sediment type, or may settle in patches brought about by feature of the larval drift. This may be distinguishable using cluster analysis and spatial autocorrelation techniques.

1. Pandalus borealis in Gulf of St Lawrence

Length frequency data and bottom temperature were collected in 1984, 1985 and 1987 during a random stratified trawl survey of 5 management areas. Following initial sorting using a non-parametric dissimilarity index, the data were examined for time-space groups using cluster analysis followed by spatial autocorrelation analysis. The data clustered consistently into five length groups (year classes) which were consistently organised spatially with respect to depth and geographical location. This is interpreted to be the result of ontogenic migration. The authors' draft text (Simard and Savard, in press) provided an extremely detailed account of the procedures and biological interpretation involved in the spatial analysis.

2. Spisula solidissima in eastern USA

Length frequency data collected by random stratified trawl survey in the Gulf of Maine were analysed on a preliminary basis by cluster analysis. The sampling strata were pre-defined independently of the biological characteristics of the surf clam distribution. Pilot results show some tendency for samples physically close together to cluster independently of the sampling strata, and to be more similar than widely separated data from the same strata. There appears to be some spatial autocorrelation in the data, and there is scope for further examination of this and a re-design of the sampling stratification.

3. Arctica islandica at Rhode Island and Martha's Vineyard

Abundance, length frequency and sediment data collected during random hydraulic dredge sampling were analysed by stepwise linear discriminant analysis. There was little or no detectable difference in size frequency between strata, but regions of high and low quahog density could be discriminated on the basis of sediment grain size. Depth appeared to play little role in the discriminant function. The report of this analysis (Fogarty, 1981) notes that these results do not necessarly mean that Arctica exercises substrate selection, as the link with particle size could be by way of some other independent factor.

D. TIME SPACE HETEROGENEITY

Biological variability and structure in a population may occur or be suspected to occur along similar lines to that observed in the previous examples, but its detection may be obscured by a variety of confounding factors. The latter include the population response to exploitation, intraspecific density, predator-prey interaction, community ecology structure, or incorrect sampling stratification. Multivariate analysis is a particularly valuable tool in searching for potentially meaningful structure in data sets from this background.

1. Barents Sea community structure from trawl surveys

The abundance of Pandalus borealis and 6 commercial species was measured by random stratified trawl surveys from 1980-1986. Standardised and log transformed data for 8 geographical sub-areas were plotted individually then analysed by PCA (7 species, 8 areas, 7 years). In the resulting projection on the first principle axis, cod and haddock were negatively associated with polar cod and capelin, suggesting a predation effect. Axis 2 is mainly associated with halibut and redfish, and to a lesser extent Pandalus. Separate projections for the 8 geographical areas showed changes in species dominance with time. Latitude, year and depth were shown to be important factors determining the distribution of species along axis 1 and 2.

2. Pandalus borealis in two Norwegian Fjords

Standardised population counts, length frequency, maturity index and dry weight were collected over two years at trawling stations in Balsfjord and Malangen. A PCA was performed on log transformed data using a correlation matrix. Groups identifiable in the projection along the first two principal axis show that Malangen data are very heterogeneous but different from those at Balsfjord, where there in turn was a systematic difference between population structure in 1985 and 1986. One station at Malangen was shown to be an outlier. Overall there appeared to be a gradient in population structure from the enclosed Balsfjord to the outermost Malangen stations near the open sea.

3. Cancer pagurus in the English Channel

Male and female length frequency data collected at three crab ports in 1985 were standardised, centred and reduced, and analysed by PCA. Variation along axis 1 was mainly due to crab size, and along axis 2 due to sex ratio. Port groupings showed both separation and overlap, because of an underlying relation between size and area on a scale broad enough to cross statistical rectangle boundaries. Sex ratio and seasonal effects were also recognised. Results were interpreted in terms of ontogenic migration and seasonal catchability and gave useful information on sampling problems.

4. Irish Sea Nephrops norvegicus

Length frequency data from a limited research vessel survey in 1988 were standardised, filtered to remove low abundance shallow water stations affected by light levels, and analysed using a dissimilarity index and complete linkage cluster analysis. Absolute abundance data showed four clusters which could be associated with different depth strata. Previous work has shown that Nephrops numbers correlate positively with an increasing proportion of silt and clay in the bottom sediments. In these 1988 data, however, differences were observed in the mean percent silt composition in the depth strata but the pattern was not systematic.

5. West of Scotland Nephrops norvegicus

A preliminary analysis was made of a large data set comprising Nephrops abundance and size, time, tide, depth and silt content, measured during 53 half hour trawl hauls in the Clyde, South Jura and South Minch. Untransformed data were analysed by PCA. The first principal component was associated with abundance, size and sediment type, the second with tidal range and depth, and a third with time and depths. Results suggest that high density populations tended to consist of smaller Nephrops on ground with low organic carbon and low median phi, and *vice versa*. The South Minch and Jura populations, though occurring at different depth, were rather similar and tended towards an intermediate to high density type. The Clyde population tended towards a low density type. This analysis needs further and more detailed study using all available variables.

6. *Pandalus borealis* in Gulf of Maine

Data from a stratified trawl survey in the Gulf of Maine were explored by PCA and canonical correlation analysis. Variables were the depth, temperature and log number of immature, male and female shrimp (distinguished by size). A plot of the high principal component scores identified a single outliner station corresponding to an unusually high abundance at low temperature and shallow depth. This point was already evident in preliminary data plots.

The canonical correlation was then calculated between the environmental variables (temperature and depth, and the population variables (immature, male and female prawn abundance). There was a significant correlation driven by abundance and temperature, and enhanced by the role of male shrimp abundance acting as suppressor variable. The parent data set is worth a detailed multivariate study.

E. GROWTH AND BODY COMPOSITION

1. *Pandalus borealis* in Balsfjord, Norway

For *Pandalus borealis* in Norway, size stratified data were available for percent weight of water, percent weight of carbon and nitrogen, total lipid, protein and ash; and weight specific energy content. Data were analysed by DFA, PCA and multiple regression analysis (MRA). Percentage data were transformed to arcsine. Data allocated to a set of *a priori* size groups deemed equivalent to age were tested by DFA and 76% of the prawns were allocated successfully to their *a priori* groups, the main determinants being total lipid, total protein and energy content. In the PCA most of the variance was accounted for by three principal axes. Protein and lipid were negatively associated (axis 1). Ash and water were positively associated with each other, but negatively associated with protein and lipid (axis 2) and water was negatively associated with energy (axis 3). Within this structure age effects could be interpreted, and in a final analysis the relative contribution of age and season were distinguished by MRA.