Fol. 41 K

<u>This report not to be quoted without prior reference to the Council</u>(*)

International Council for the Exploration of the Sea

C.M. 1989/K:38
Ref.: D

Conseil International pour l'Exploration de la Mer

REPORT OF THE WORKSHOP ON SPATIAL STATISTICAL TECHNIQUES

16-19 May 1989, Brest, France.

Co-conveners:

Gerard Y. Conan
Dimitri Stolyarenko

This document is a report of a Workshop of the International Council for the Exploration of the Sea and does not necessarily represent the views of the Council. Therefore, it should not be quoted without consultation with the General Secretary.

*General Secretary
ICES
Palaegade 2-4
DK-1261 Copenhagen K

TABLE OF CONTENTS

---ooOoo---

1 SUMMARY

A workshop on spatial statistical techniques met for 4 days in Brest France. The purpose was to evaluate statistical techniques recently introduced in fisheries research for mapping and assessing the resources after taking into account their spatial dispersion patterns. A knowledge of the location of the sampling units is required. These techniques present interesting perspectives for optimising the sampling design during sampling operations, for correcting possible biases in sampling schemes that cannot be randomly designed, and for analyzing the spatial dispersion of marine organisms and understanding the factors that determine these dispersions.

ABREGE

Un atelier s'est reuni a Brest, France pour etudier durant quatre jours les techniques de statistiques spatiales. Ces techniques ont ete recemment introduites en recherche halieutique, pour cartographier et evaluer les ressources en tenant compte de leurs patrons de repartition spatiale. La connaissance de la position des unites d'echantillonnage est requise. Ces techniques presentent des perspectives interessantes pour optimiser le plan d'echantillonnage durant les operations, pour corriger les biais qui risquent d'etre introduits lorsque le protocole d'echantillonnage ne peut etre aleatoire, et pour analyser la repartition spatiale des organismes marins en interpretant les facteurs susceptibles d'etre a l'origine de ces repartitions.

2 INTRODUCTION

This workshop was voluntarily not organised in the form of a formal assessment meeting. The intent was not to take decisions and to make formal recommendations to the Council concerning the opportunity to modify traditional survey methodologies after reaching a consensus. The intent was to create a constructive forum for exchange of ideas among specialists coming from very different backgrounds, using very different techniques, but sharing a common interest in incorporating spatial patterns into statistical analysis and modelling of survey data.

Therefore, after consulting several participants, and in order to avoid the limitation of consensus generalities, this report was prepared in the form of minutes of a debate. Several rapporteurs provided their input on each of the presentations made, ideas were later presented to the chairpersons by mail. The chairpersons attempted to record these ideas in their free form and by no means imply that any of these ideas were accepted by

all participants. It was felt that it was unpractical to attempt to have this report approved by all participants due to the preliminary nature of the discussions and the lack of common background for the debates.
The recommendations presented at the end of this document were written and presented by participants during the meeting. These recommendations are of a very general nature and summarize the views of the participants concerning the need for reconducting the workshop in the form of a scientific forum and for developing tools incorporating spatial effects into assessment methodologies. The workshop did not attempt to force the endorsement of any specific methodology and it was felt that this was not part of the mandate of the meeting.

2.1 Participation

Mr. S.A. Akenhead, Dept of Fisheries & Oceans, NW Atlantic Fisheries Centre, P.O. Box 5667. St John's, Nfld A1C 5X1, Canada.

Mr K.P. Andersen, Danmarks Fiskeri-og Havundersogelser Charlottenlund Slot, 2920 Charlottenlund, Denmark.

Dr N. Bailey, Marine Laboratory, P.O. Box 101, Victoria Road, Aberdeen AB9 8DB, U.K.

M. P. Berthou, IFREMER, Centre de Brest, BP 70, 29263 Plouzane, France.

Ms A. M. Caramelo, INIP, Avenida de Brasilia, 1400 Lisbon, Portugal.

Dr B. Carbonell, Institut fur Meereskunde der Universitat Hamburg, Troplowitzstrasse 7, 2000 Hamburg 54, Federal Republic of Germany.

Ms M. de F. Cardador, INIP, Avenida de Brasilia, 1400 Lisbon, Portugal.

M. R. Chevalier, IFREMER, rue de l'Ile d'Yeu, BP 1049, 44037 Nantes Cedex, France.

Dr G.Y. Conan, Laboratoire de Recherche en Biologie marine, Universite de Moncton, Moncton, N.B. E1A 3E9, Canada.

M. D. Corre, ORSTOM, Centre IFREMER, 29273 Brest, France.

Dr C.D. Darby, Fisheries Laboratory, Lowestoft NR33 OHT, Suffolk, U.K.

M. H. Demarcq. ORSTOM, Centre de Recherches Oceanographiques de

Dakar-Thiaroye, BP 2241, Dakar, Senegal.

M. R. Dijkema, Netherlands Institute for Fishery Investigations, P.B. 77, 4400 A B Yerseke, Netherlands.

Dr. K.G. Foote, Institute of Marine Research, P.O. Box 1870, Nordnes, 5024 Bergen, Norway.

M. D. Gerdeaux, Institut de Limnologie, B.P. 511, 74203 Thonon-les-Bains, France.

M.M. Girardin, Centre National de Recherches Oceanographiques et des Peches, BP 22 Nouhadhibou, Mauritanie.

M. F. Gohin, IFREMER, Centre de Brest, BP 70, 29263, Plouzane, France.

Mr G. Goussev, Ministry of fisheries of the USSR, Moscow, 12 Rozhdestvensky Bld, Moscow, USSR.

Dr J.P. Hillis, Dept. of the Marine Fisheries Res. Centre, Abbotstown, Dublin 15, Ireland.

Mr H. Hinze, Alfred-Wegener-Institut f. Polarforshung, Columbusstrasse, 2850 Bremerhaven, Federal Republic of Germany.

M. Inejih Cheikh Abdallahi, Centre National de Recherches Oceanographiques et des Peches, Nouadhibou BP 22, Mauritanie.

M. E. Josse, Centre ORSTOM/IFREMER. BP 70, 29263 Plouzane Cedex, France.

Mr M.C.S. Kingsley, Dept of Fisheries & Oceans, Inst. Maurice Lamontagne, 850, rte de la Mer, P.O. Box 1000, Mont Joli, Quebec G5H 3Z4, Canada.

Mr G. Lavoie Universite de Montreal, C.P. 6128, Succursale "A", Montreal (Quebec) H3C 3J7, Canada.

M. J.J. Lechauve, ORSTOM, Centre de Brest, BP 70, 29263 Plouzane, France.

Mr J. Modin, Institute of Marine Research, Box 4, 453 00 Lysekil, Sweden.

Dr M. Moriyasu, Dept. of Fisheries & Oceans, P.O. Box 5030, Moncton, N.B. E1C 9B6, Canada.

Mr A. Nikolajsen, Fiskerirannsoknarstovan, Debessartrod, Fr-100 Torshavn, Faroe Islands, Denmark.

M. P. Petigas, Centre de Geostatistique, 35, rue St Honore, 77305

Fontainebleau, France.

Dr T. Polacheck, Northeast Fisheries Center, NMFS/NOAA, Woods Hole, Mass. 02543. USA.

M. J.C. Poulard, IFREMER, 8, rue Francois Toullec, 56100 Lorient, France.

M. D. Renard, Centre De Geostatistique, 35 rue St Honore, 77305 Fontainebleau, France.

Dr J. Rivoirard, Centre de Geostatistique, 35 rue St Honore, 77305 Fontainebleau, France.

Dr Y. Simard, Dept of Fisheries & Oceans, Inst. Maurice-Lamontagne, 850, rte de la Mer, P.O. Box 1000. Mont Joli, Quebec G5H 3Z4, Canada.

Dr Sorokin, VNIRO, 17 Verkhne Krasnoselskaya, Moscow B-107140, USSR.

Mr T.K. Stokes, Fisheries Laboratory, Lowestoft NR33 OHT, Suffolk, U.K.

Dr D. Stolyarenko, VNIRO, 17 Verkhne Krasnoselskaya, Moscow B-107140, USSR.

Mr. K. Sunnana, Institute of Marine Research, P.O. Box 1870/72, Nordnes, 5024 Bergen, Norway.

Mr S. Tveite, Flodevigen Biological Station, 4817 HIS, Norway.

Mr J.H. Volstad, Institute of Marine Research, P.O. Box 1870/72, Nordnes, 5024 Bergen, Norway.

Mr E. Wade, Dept of Fisheries and Oceans, P.O. Box 5030, Moncton, N.B. E1C 9B6. Canada.

Dr R. M. West, Orchard House, Lower Park Road, Braunston, Devon, U.K.

2.2 Mandate

As stated by the Council resolution C.Res.1988/2:19, taken during the 1988 statutory meeting:

A Workshop on Spatial Statistical Techniques will be established under the chairmanship of Dr G. Conan (Canada) and Dr D. Stolyarenko (USSR) and will meet in Brest for 4 days in May 1989 at national expense to:

a) review the theoretical basis for applying geostatistical techniques to the study of spatial dynamics,

b) review their application to shellfish populations,

c) test the robustness of different geostatistical approaches for analysing a variety of observed and simulated population survey data, and

d) compare and exchange associated software.

2.3 Origin, aims and structure of the workshop

At the 1988 ICES statutory meeting in Bergen, methodologies dealing with the analysis of spatial distributions of fisheries resources, and with automated mapping and assessing, generated great interest.

Traditionally, spatial distributions are ignored by Fisheries Science. For the purpose of assessing a resource, a general practice consists in randomly distributing the sampling units in order to avoid any bias created by non-independence.

The workshop was set up to consider the alternative solution of studying the spatial non-independence of the sampling units, modelling this non-independence, and integrating it into the estimates. Instead of avoiding bias, this approach opens the perspective of dealing with bias. This may allow for a greater flexibility in 1) optimizing the sampling design and, 2) for processing data for which the experimentator has a restricted control or no control on sampling design. Examples are acoustic surveys in the first instance, and statistics on catch from a commercial fishing fleet in the second instance.

The workshop was also to consider the applications of automated mapping of a resource for incorporating patchiness in the estimation of the economically harvestable portion of this resource. The harvestable portion of a resource frequently depends on the spatial repartition of this resource into high density patches.

The workshop was also to consider possible generalizations of spatial models and techniques to populations of marine organisms, particularly sessile or semi sessile invertebrates, and the ecological significance of certain spatial patterns most frequently encountered.

In order to generate new challenging ideas, the workshop was to be informal, international and multidisciplinary. Experts were to be brought from backgrounds as varied as Fisheries Science, Marine Ecology, Oceanography, Statistics, Informatics, and Mining Resource Surveying (Geostatistics).

3 SCOPE OF SPATIAL STATISTICAL TECHNIQUES

Spatial statistical techniques derive from at least four schools of thinking:

a) Ecologists conscious that the spatial dispersion of organisms is usually not "random", but displays patterns which usually prevent sampling distributions to be modelled by a Poisson probability distribution (Elliot, 1971). The dispersion is said to be "patchy" when the ratio of variance to mean is significantly greater than 1, regular when it is significantly smaller than 1, and random when it is not significantly different from 1. This type of approach usually deals with spatial patterns existing at a scale smaller than the area covered by a sampling unit. Information on the location of the samples is not required. Probability distributions such as binomials (positive and negative), Poisson, lognormal, gamma, delta are used for modelling the sampling distributions and fitting confidence limits to the estimate of the mean.

b) Geographers considering the geographic distribution of some quality or phenomenon within predefined geographic entities, countries or states and questionning whether the presence of some quality in a country makes its presence in neighbouring countries more or less likely (Cliff & Ord, 1973). This approach has been extended to ecology, it consists in testing for the presence of spatial "autocorrelation" i.e. non independence between geographic entities, eventually sampling units taken during a survey, and in incorporating autocorrelation concepts in statistical tools such as analysis of variance or regressions.

c) Mining industry engineers considering the problem of optimizing predictions of abundance of a resource in an unsampled location, when sampling units have been taken at known locations in the neighbourhood, usually not randomly, but with the aim of maximizing profit. This approach has diverged towards optimized automated generation of charts of a resource, and optimized assessment of a resource over a given region, with assorted confidence limits. In this approach non independence of sampling units taken at close range is an undisputed fact which does not need to be tested, but needs to be modelled. The model is incorporated in the estimates.

d) Mathematicians and computer scientists developing models capable of generating forgeries of landscapes by "fractal" geometry (Mandelbrot in Peitgen & Saupe eds, 1988). This approach allows to model the complexity of spatial patterns in nature, and to fit indices of complexity, the "fractal dimensions". The models have sometimes been incorporated in assessment techniques for estimating the abundance of a resource within associated confidence limits.
Most assessment techniques used in fisheries science fit within the philosophy of school of thinking (a).

4 PROCEEDINGS

The presentations and topics discussed are presented in a chronological order.

4.1 **Opening by G.Y. Conan:** Mandate of the workshop and growing interest for spatial statistic in fisheries biology. Kriging is now officially recognized and routinely used for mapping and assessing snow crab resources in the Gulf of St Lawrence, Canada.

4.2 Introductory overview by Pierre Legendre: Analysis of spatial patterns in ecology.

The spatial heterogeneity of populations and communities plays a central role in ecological theory describing succession, adaptation, maintenance of species diversity, community stability, competition, predator-prey interactions, parasitism, epidemics and other catastrophes, ergoclines. The presentation showed how the spatial structure of biological populations and communities can be studied. Many of ht basic statistical methods used in ecology are impaired by autocorrelated data. Most if not all ecological data is autocorrelated. In order to remain valid, most statistical tests must be adapted in the presence of spatial autocorrelation.

Methods presently available for analyzing the spatial structure of biological populations were described and illustrated in the case of vegetation data. Various methods allow to test for the presence of spatial autocorrelation in the data: all directional and two directional spatial correlograms, two dimensional spectral analysis, and the multivariate Mantel test and Mantel correlogram. Other methods are descriptive, such as the univariate variogram, and the multivariate methods of clustering with spatial contiguity constraint, the partial Mantel test, presented as a way of studying causal models that include space as an explanatory variable. Various methods allow to map ecological variables: interpolation, trend surface analysis, kriging. Reference: Legendre & Fortin, in press.

4.3 Introduction by G.Y. Conan: Basic concepts of point kriging, block kriging, non linear kriging used in selected case studies.

4.4 Case studies presented by Y. Simard, D. Renard, P. Petitgas, M. Moriyasu respectively on pandalid shrimp, clams, hake, and snow crab.

The software packages Surfer, Bluepack, and Gulfkrig have been used respectively for dealing with 1) shrimp, 2) clams and hake, and 3), snow crab stocks. In the first instance the aim of the study was mainly to map the dispersion of shrimp through time and space, these appear to be very patchy and unstable, a considerable challenge for optimizing the design of a standard stratified random assessment. The second case delt with mapping as well as estimating a stocks of clams, and of hake using the theory of intrinsic random function of order k; the resource is extremely patchy and very localized peaks in abundance depart from the general spatial structure of the stock, this can be satisfactorily modelled by IRFK. The last case uses linear kriging for mapping and assessing a very patchy stock of snow crab, confidence limits were provided with the estimates of global abundance and a comparison was made with standard estimates using the arithmetic mean; the kriged averages tend to be smaller than the arithmetic mean, possibly due to an involuntary bias in the sampling procedure; the kriged variance tends to be always larger than the variance of the arithmetic mean, in the presence of autocorrelation the latter would tend to provide overoptimistic confidence limits.

The following elements of discussion were reported by the rapporteur P. Petigas:

a) Estimates and assumptions. A linear kriging estimate does not need any distributional hypothesis, but requires a model of the spatial structure. In standard statistics, a distributional model is needed. For a global estimation, it has been agreed that if the spatial structure is accounted for, this lowers the estimation variance. Concerns were raised about the consequences of a poor fit of the spatial model for global estimates. It was stressed that assuming no spatial structure is actually implicitly the same as fitting a 0 covariance model. It was felt that comparisons were required.

b) Spatial structure, global behavior and local effects. Inhomogeneity of the variance due to non independence of the variance to the mean have been found to create difficulties for modelling an overall spatial structure due to local effects.

c) Limits of spatial domain. A global estimate is defined within a given spatial domain (volume or area). The search for an appropriate method for handling zero values has raised considerable interest. Several techniques have been proposed: simulation of a zero contour, decision to set a threshold on the variance of estimation after mapping. No general recommendation was made. Fortunately global estimates of densities do not appear to vary considerably when calculated on different domains, there is not yet sufficient experience as to how the variance behaves.

Interaction of stratification with spatial structure was considerably discussed. Visualization of maps and changes in spatial distributions through time was very well received with the example of pandalid shrimp data. Maps were considered a way to introduce environmental information in fisheries biology and management.

4.5 Presentation of software and hardware facitilies available at IFREMER by F. Leverge.

IFREMER Brest has considerably invested in the development of their own GIS and image processing systems based on custom designed software. The kriging technology is used in satellite surveys of oceanographic parameters. An elaborate computer network based on main frames, SUN workstations, and personal computers provide easy access to data bases and software utilities throughout IFREMER campuses, including ships, Tahiti and New Caledonia.

4.6 General introduction by G. Conan. Sampling vs assessing.

Sampling designs in fisheries research are always supposed to be perfect random or stratified random designs. Unfortunately many factors can cause a design not to be random. The structure of certain instruments also prohibits random sampling (aerial surveys, acoustic surveys). In the design of strata, preliminary information is required on the spatial structure of the resource, frequently a depth contour is substituted, although it might not even be correlated with the actual stock structure. Can spatial statistics methods obliviate the need for random sampling ? in which case what would be an optimal sampling design ?

4.7 General overview by M. Kingsley: Sampling for resource evaluation, in the context of standard statistics.

The following is a edited version of J.Volstad's rapporteur's report. Sampling from populations which are continuous and structured was discussed with emphasis on systematic vs random sampling designs. For most aerial transect surveys, a systematic design is used.

Systematic sampling has operational advantages: it is easy to move in straight lines, it is easy for navigation, it provides good distributional information, it is cost efficient.

Systematic sampling is precise for non-periodic populations, and the correlogram is everywhere upward concave, it is more precise than random sampling. However, unbiased estimates of the error variance are not available. The process of systematic sampling with a random start is unbiased.

Random samples are "insured" against unknown structures of the populations sampled, and error variance can be estimated. If extra survey resources are made available, the survey can easily be continued without causing problems for the estimation procedure.

Ensuing discussion:

a)Variograms may provide a logical basis for stratification. Variograms change according to sampling unit size, i.e. grouping of observations along the transect will affect the shape of the variogram, but he theory of these changes is well understood and can be accounted for.

b)For acoustic surveys, in a random design more information is lost when transects are close than is gained when they are further apart. In general, the workshop agreed that systematic transects are the most efficient in terms of precision.

c)No overall best sampling design exist in terms of standard statistics. The estimation of resources may require a different design than the description of their spatial structure by mapping.

4.8 Optimisation of sampling design in a geostatistical context, by J.Rivoirard and D.Renard.

The following is an edited version of J. Volstad's rapporteur's report. The objective is to map the density of biomass with a maximum precision, and with a representation of real variability.

Transitive theory provides unbiased estimates with corresponding variance for the surface. This approach is useful for global estimation problems, not for local estimates. No assumptions are made about stationarity. However, the observations must be from a regular grid.

In the usual kriging approach the variogram permits to estimate the error variance of the linear estimates for the stationary case. In the non stationary case, generalized covariances of order k should be used. When estimating global biomass density, the variance can be obtained for any sampling design. Local estimates of biomass, with associated variance can be obtained. Kriging provides minimum variance estimates by using optimum weights.

Non linear estimates may be obtained from simulations.

Problems arise when it is required to estimate the probability of exceeding a threshold (this can occur in pollution problems). Kriging estimates are not optimum for this case.

Ensuing discussion:
The participants discussed but did not reach a consensus as to whether kriging estimates were unbiased. G. Conan mentioned that in his practical experience of processing fisheries data, the kriging averages frequently tended to be lower than their arithmetic mean counterpart and attributed this to the fact that "random samples" were frequently involuntarily biased towards high density locations. The method of optimising the weights built into kriging tends to correct such a bias. He also mentioned that the kriging variance tends to be higher than the variance of the arithmetic mean, due to the same effect, the variance of the mean of such a slightly biased sample will be smaller than the variance of the mean of a random sample.

G. Conan mentioned that in the absence of covariance effects, i.e for a flat variogram, the kriging average and the arithmetic mean, the kriging variance and the variance of the mean are exactly identical. Several participants held that kriging averages were unbiased, as the arithmetic means for random samples, but no consensus was reached, the main contention being the fit of the variogram.

4.9 Analysis of variance for autocorrelated regional data, by P. Legendre.

The following is an edited version of J.Volstad's rapporteur's report. Usual ANOVA statistics are biased if the data are spatially correlated within geographic regions. Positive autocorrelation reduces within group variability, thus artificially increasing the relative amount of between-group variance. Negative autocorrelation may produce the opposite effect. This can be viewed as a loss of an unknown number of degrees of freedom. A method was presented which in certain cases overcomes this problem. It is based on the computation of pooled within-group sums of squares for sampled permutation of internally connected areas on a map. A computer program for the analysis is available at no cost.

Ensuing discussion:
It was mentioned that most fisheries data are likely to be spatially autocorrelated and that one should be cautious when using ANOVAS for the analysis of CPUE's.

4.10 **General introduction on the potential uses of spatial statistics for processing transect data, by G.Conan.**

There are various instances for which data is necessarily collected along transects: acoustic surveys, aerial surveys, satellite surveys. Such data are frequently collected for shellfish biomass or associated parameters such as effort, or surface water temperatures. Can spatial statistics provide an answer to such difficult cases when standard statistics do not provide a comprehensive approach ?

4.11 **a) Problems encountered in sampling and integrating acoustic signals for mapping and assessing a resource, by K. Foot.**
b) Examples of echo survey in two lakes, Annecy and St Croix, use of geostatistics, by D. Gerdeaux.
c) Mapping and assessing a school of herring using the kriging technique, by E. Wade and G. Conan.
d) Geostatistics and teledetection, by F. Gohin.

The following is an edited version of N. Bailey's rapporteur's report. The application of spatial statistics to two types of data collected by remote sensing was considered. The first type, acoustic survey data (used in mapping, abundance estimates and behavioral studies) is typically collected along transects across the area of interest. The echo signal provides considerable quantities of detailed information along the transect but nothing in between transects. The mapping of hydrographic features provided the second example where spatial statistics were used. The method may be equally applicable to other cases such as aerial surveys.

Three presentations on acoustic surveys were made; the first, by K. Foote, outlined the problems encountered in sampling and integrating acoustic signals for mapping and assessing abundance of a resource. There has been for some time a realization that sampling strategy and methods of integrating the data for assessing the abundance of the resource, were not efficient compared to the high resolution of the tool. It is only recently that discussions have begun as to how these difficulties could be resolved.

The problems include the choice of the area to cover, the sampling strategy (random, systematic, etc.), the difficulty of carrying out preliminary surveys when the resource is highly mobile, the presence of various species and the structure of the population which demand additional information from trawling. Integration of the data is often achieved by linear interpolation and by taking the arithmetic mean. Autocorrelation problems are known to exist but the effects are mostly ignored.

Various methods of data processing were proposed during the ensuing discussion, including spline and kriging.

Two case studies of kriging applied to acoustic surveys were presented , one by D. Gerdeaux, on the application of geostatistics to echointegration in freshwater, and one by E. Wade and G. Conan on mapping and assessing a school of herring using kriging.

Various approaches were assayed for improving the definition of the variogram. In one instance additional transects were taken to supplement the main grid and better define the shape of the variogram close to its origin (nugget effect). In the other instance the total number of unit echos (pings) along the transect were used as individual sampling units.In both studies the area covered was partitioned into subzones for which separate variograms were fitted.

The lake study showed that where the distribution of fish was even, the stock estimate obtained by arithmetic mean and geostatistics were similar, however, when there was heterogeneity, different estimates were produced. Similar differences between arithmetic mean estimates and geostatistics estimates were also found in the analysis of the school of herring.

The importance of ensuring a complete coverage by adequately designing the transect pattern, and to identify trends was emphasized. In the lake survey there was a higher abundance around the banks and a lower abundance towards the centre. In the herring survey, a higher abundance towards the north.

The lower number of points towards the edges of the survey area creates a higher variance for local estimates. In order to avoid the associated lack of precision, the transects should be drawn beyond the limits of the patch. This will also ensure a better definition of the patch.

The shape of the variogram produced from acoustic surveys reflects the schooling behavior of the resource. In the hake study, a drop in semivariance at a distance (lag) of about 4 Km was attributed to a structure made out of subpatches some 4 Km in diameter.

A fractal model was used to describe the variogram of the school of herring and generalizations were proposed on the significance of fractal dimensions calculated from variograms and on case studies in which the variance may never stabilize what ever the lag.

The study on herring data addressed the problem of irregular shaped areas frequently surveyed by acoustics and used a specifically designed software instead of the more general discretizing procedure used by many commercially available geostatistical packages. The software also permitted constraints to be imposed which took into account a coast line both for mapping and for resource estimation.

The talk by F. Gohin illustrated the use of geostatistics in mapping environmental features such as sea surface temperatures. An interesting development was the simultaneous use of data from various sources, in this case, temperatures recorded by ships and by remote sensing from satellites. Structural analysis using each type of data separately was used to compare results and construct maps from the combined information.

ensuing discussion:

a) The definition of the type of model required for fitting the variogram and the quality of the fits were discussed as well as the need to remove trends <u>prior</u> to fitting the variograms. No consensus was reached. The most of geostatistical techniques consider that, since the data is autocorrelated it is incorrect to remove trends prior to fitting a variogram, both the variogram and the trends should be fitted simultaneously.

b) The subdivision of the area into subzones in which separate variograms were fitted was considered by some participants as a process equivalent to post stratification in which the objectivity of the partitions should be questioned.

c) The opportunity to introduce trends in global estimates by kriging was discussed. G.Conan indicated that in his experience of data processing using kriging, global estimates with or without trends tended to be very similar as long as the sampling units were adequately dispersed over the whole area assessed; differences became important when extrapolations were made over an unsampled area.

d) Some participants suggested testing for significance of the difference between arithmetic mean and kriged average by using a t test based on their respective variances. Others considered that this was meaningless due to the fact that the method and not the data sets differed.

e) The adequacy of fractal models in geostatistics was discussed but no consensus was reached. Some of the contenders held that a confusion was being made between fractal structures and "self similar" structures when it was stated that a fractal structure displayed similar complexity whatever the scale considered. However these participants were referred to R.F. Voss's (<u>in</u> Peitgen & Saupe eds, 1988) definition backed up by Mandelbrot (the developer of fractal theory): *"Mandelbrot's fractal geometry provides both a description and a mathematical model for many of the seemingly complex forms found in nature. Shapes such as coastlines, mountains and clouds are not easily described by traditional Euclidian geometry. Nevertheless, they often possess a remarkable simplifying invariance under changes of magnification. This statistical self-similarity is the essential quality of fractals in nature. It may be qualified by a fractal dimension, a number that agrees with our intuitive notion of dimension but need not be an integer."*. G. Conan comments that the differences of vue mainly rely on semantics. A fractal variogram as per Voss's definition is simply represented as a linear model on a log log plot. An interesting feature is that it has no range which implies that the variance never stabilized whatever the distance between sampling units. A simple minded way of visualizing this type of spatial pattern is that patches would have a structure similar to clouds.

e) Acoustic data may have a strong temporal component which is not normally considered in usual geostatistics. The risk of recording the same fish several times if the spatial structure is not stable was also raised, but it was pointed out that this problem exists in any method and is relevant to the sampling tool, not to the methodology of data processing.

f) The issue of separating the error into components was addressed. With modern navigation systems, positioning was not considered to be a major source of error. The nugget effect is not only explained by instrumental error. It is difficult to dissociate sampling error from white noise error.

g) There was a general agreement that the application of spatial statistical techniques to remote sensing data, and particularly to acoustic data would extend the usefulness of these techniques and was a worthwhile future research area. Comparative studies are desirable for testing the reliability of the different spatial statistics methods in producing maps of dispersion and stock abundance estimates.

4.12 **Ecological developments in spatial statistics.**

Two talks were given under this heading, Scott Akenhead assumed the tasks of rapporteur, an edited version of his report follows.

a) Geographic segregation of biological categories in snow crab, by M.Moriyasu. The spatial dispersion patterns of distinct life history categories of individuals in a population of snow crab were analysed in the Gulf of St Lawrence. The spatial dispersions are disjunct but interact, some are positively intercorrelated, other are negatively intercorrelated. It appears that some of these exclusions and attractions are related to behavioral reproductive patterns. A revealing contribution by G. Conan used correspondence analysis to define synthetic variables (Principal axes) which could be related to ontogeny and maturity. The composite values obtained from the projections on the principal axes of each of the multivariate observations were mapped and displayed a spatial structure of the ontogenic axis. Complex behavioral geographic exclusion processes linked with mating and reproduction of snow crab may create these patterns.

b) Mapping multivariate data: constrained clustering, by P. Legendre. A biogeographic type of analysis was proposed. The technique ensures that contiguous geographic areas result from the clustering. A software package (COCOPAR) is available. Examples related to the analysis of fish communities and forestry provided good demonstrations of the ecological value of this new tool.

4.13 **Spatial statistical tools for mapping and assessing.**

G.Conan suggested that a feeling for spatial statistics theory would be opportune once the participants all shared a feeling for the approaches used for analysing the case studies and the sets of their own data on the software demonstration packages. During this session "user friendly" presentations of the available techniques for mapping and assessing were to be made. Emphasis was not set on mathematical developments but on the underlying concepts of these techniques.

4.13.1 Kriging fundamentals, a review, by D. Renard and J. Rivoirard.

S. Akenhead acted as rapporteur and provided the elements for the following report. Kriging consists in allocating optimised weights to each value associated with a sampling unit. These optimised weights depend on the variance associated with the distances, between sampling units locations, and between the sampling units and the location for which an estimation is being made. The variances are simply read from a variogram previously fitted to the observed values of variance associated with a distance between pairs of points. The sum of the weighted values is the kriging average. A kriging variance can be calculated as a function of these weights. The ability to add subjective constraints allows to bound the body being considered.

4.13.2 Spline approximation, by D. Stolyarenko.

The following is an edited version of R. Chevalier's rapporteur's report.The spline approximation method can be applied to complete trawl and acoustic surveys and used for designing and conducting surveys. The method provides the opportunity to merge the measurements with self-calibration. This is based on a cross validation algorithm.

The spline approximation method uses the relative position of the measurements. The closest the points are located, the stronger the sampling unit measurements are correlated. The smaller a distance from a sampling unit to the point to be estimated, the stronger the influence of the value of this sampling unit on the value of the density in the reconstructed point. This property is shared with kriging and it allows non-random survey designs, for instance adaptative sampling with replicate measures taken in high density locations.

The meaning of the term "distance" can be redefined, non Euclidean distances can be used. To the simplest anisotropy can be taken into account in the distance, and this anisotropy can vary as a function of the location of the sampling unit. Depth d can be incorporated as part of the distance r in the form: $r=(x^2+y^2+k.d^2)^{.5}$.

Measurements are associated with a "noise", as in a regression this noise must be smoothed out. In data smoothing the challenge is to balance the smoothing of the density with fidelity to the measurements. An interpolating function which passes through all the measurements overestimates the accuracy of the measurements and is inadequate when replicated measurements including noise are entered.

A linear function is inadequate at the other extreme, it underestimates the accuracy of the measurements. In this respect spline approximation differs from kriging, cross validation is used for testing locally the goodness of fit: each real measure is compared with its predicted value, smoothness is adjusted to minimize prediction error. The maximization is made with respect to σ (the standard deviation) and the vector of parameters $=(_1, _2, \ldots _i)$, m, in a first approach using generalized least squares. Further the maximization can be achieved by a simplex method to obtain the function of correlation.

The software includes a geographic data base of depths in order to provide the possibility to include depth in the calculations when these are not measured at the time of the survey. The technique has been applied successfully to shrimp, crab, scallop, seaweed, cod and redfish stocks.

4.13.3 The Linear Generalized Model, by T. Polacheck, J.H. Volstad, and R.M West

The method can be applied to marine surveys of stock abundance whatever the location of the sampling stations. Rather than modelling the fish counts which have a skewed distribution, the method will model S, the square root of the fish counts:

$S= y_1 + X_2.y_2 + X_3.y_3 + X_i^i.y_i + \epsilon$

where

$X_.$ represents a variable

$y_.$ represents a parameter to be fitted

Two residuals at sample stations a distance r apart are positively correlated, the correlation between them diminishing or increasing when r varies. Parameters are estimated by maximizing a likelihood function. The model can be used for prediction purposes at positions other than sample stations. The model calculates simultaneously a drift and a variogram, but contrarily to kriging that makes no assumptions on the distribution, the model supposes a multinormal distribution.

The model was applied to Chlamys islandica by Volstad and West (1989) and fitted well the data.

4.13.4 Non Parametric methods used for processing oceanography data and mapping fish, by S. Akenhead.

Using information on various accessory variables such as bottom type, temperature, abundance of another species can improve abundance estimates. Evans, Rice, and Akenhead have successfully used accessory variables for improving fish estimates. This technique is called an "external drift field" by geostatisticians.
However difficulties are encountered in the existence of covariances between the residual error in the accessory variable and the residual error in the geostatistical estimate of the abundance of

fish. Intrinsic Random Functions of Order K allow to deal with this difficulty, strictly local considerations remove the covariates effects, accounting for the variance of residuals.

The shape of the sampling distributions as well as the variances may vary locally. Non parametric methods may allow to deal with this problem.

4.13.5 Ensuing common discussion.

Rapporteur: S. Akenhead

a) The statistical purity of using the data twice by first estimating a variogram and later obtaining a kriged estimate over the surface was questioned.

b) The group agreed that kriging could correct for preferential sampling in the process of providing global estimates of a resource. The preferential sampling represented by commercial fish catch data was discussed and the idea of trying to combine commercial and research vessel data through geostatistical techniques was found most interesting.

c) Kriging can be used 1) for integrated estimates of biomass over a region, 2) for mapping regionalized estimates calculated on a high resolution grid.

d) Adaptative sampling during the survey is possible and may result in substantial gains in efficiency.

e) Utilizing more than spatial location for defining the distances between points could be very useful e.g. depth, temperature, bottom type. Extensive fields of such factors may generate problems of covariance with the residuals of the desired estimates,. Using the "Intrinsic Random Functions of order k" development for geostatistics models is a solution to the problem, although mechanistic explanations are not part of IRF technique.

f) The "step function" approximation to spatial modelling, that is to say stratified random analysis will usually be better (and never worse) for providing a global estimation (in terms of bias and precision) than a simple mean and variance. It needs to be determined whether the "smooth function" approximation, that is to say, kriging, spline, etc. provide only an improvement upon the step function approximation. Simulation studies of this problem are called for. It may be that important spatial autocorrelation exist within strata, leading to underestimates of within strata variance, and seemingly increased precision of the global mean.

e) Splines may be equivalent to kriging in a stationary case. Some participants held the opinion that splines implied the presence of a drift while others believed that splines were equivalent to kriging with a fixed model as variogram. There appears to be different uses of the word "spline" referring to approaches differing from the technique presented.

4.14 Comparison of the performance of the diverse spatial statistical techniques used for mapping and for global estimates.

Several participants expressed the need for comparing mapping and assessment techniques by simulation. G.Conan asked P. Legendre to present a paper illustrating such a comparison as a key note closure for the meeting: A comparison of kriging and other surface estimation techniques. (Reported by G. Lavoie).

The objective was to compare the accuracy of kriging algorithms and other surface estimation methods in reconstructing known surfaces from data drawn (sampled) from these surfaces by simulation.

Three sampling designs were applied to four different surfaces: uniform random distribution, clumped distribution, regular distribution on a lattice, applied to, a single fat bump, several fat bumps, single thin bump, several thin bumps. Each simulation was repeated 50 times. The surface fitting techniques were, maximum likelihood, kriging, trend surface interpolation, nearest neighbour polygons, inverse square distance weighting and averaging.

The goodness of fit was measured as the square root of the sum of squares of the deviations. None of the methods always performed better than all others, their power of reproducing the best part of the spatial dispersion of the originating variable depended upon the type of simulation.

5 CASE STUDY

A case study was proposed to the participants, in order to try out the different methods available. The set of data is small enough to allow easy calculations which can be cross checked manually if necessary.

The data consist of 29 data points representing density of herring eggs on a benthic spawning site in the Gulf of St Lawrence (Canada), it was provided by Dr Shoukry Messieh, Bedford Institute of Oceanography, Canada (Messieh S., 1988; Messieh S.N., and H. Rosenthal, 1989).

North Km	East Km	Egg density (x1000/m2)
1.703	0.778	15.46
1.333	0.8148	12.11
1.333	0.3889	7.93
0.9630	0.4074	9.12
0.1852	0.463	1.51
0.0	0.0	0.12
0.1852	0.0	1.63
0.3704	0.0	2.52
0.5556	0.0	0.78

0.7408	0.0	1.46
0.9815	0.2963	0.23
0.6852	0.240	3.97
0.5556	0.185	1.16
0.1852	0.185	0.06
0.0	0.1852	0.61
0.0	0.3889	1.48
0.3704	0.2222	0.80
1.333	0.1852	0.85
1.333	0.5741	11.29
1.703	0.5741	7.63
1.889	0.7778	8.68
2.092	0.9630	6.51
1.889	1.148	0.67
1.518	0.778	6.35
1.889	0.963	2.39
1.703	0.9630	2.21
0.7408	0.3889	1.68
0.5556	0.3889	0.91
0.3704	0.3889	0.14

Five participants responded to the request of processing this data and provide material to be included in this report.

a) Tom Polacheck: Testing for the presence of spatial covariance and trends by GLM methods:

The data was fitted to the following model:

$$Z_i = b_0 + b_1 X_i + b_2 Y_i + e_i \qquad (1)$$

where:

Z_i = *egg density*
X_i = *North coordinate*
Y_i = *East coordinate*
e_i = *the residuals*

The residuals are assumed to have variance ^{2} and to be spatially correlated. Two functions were explored for their relationship between the correlation of the residuals and the physical distance between samples:

$$corr\ (e_i,\ e_j) = exp\ (-r_{ij}/L) \qquad (2)$$

and

$$corr\ (e_i,\ e_j) = exp\ (-r_{i,i}/L)_2 \qquad (3)$$

where:
r_{ij} = *the distance between samples i and j*
L = *a constant*

The author also considered using the egg densities without a transformation and will do a square root transformation.

The main results were consistent across both sets of correlation functions and transformations. The best fit significant model in all cases was the model that included only a linear term for the north coordinate (b_1) without any spatial covariance terms.

If no linear terms are included in the model (i.e only a simple mean is included), the spatial cvariance term (L) for either transformation is significant for function 1 but is not significant for function 2. In all cases the likelihood value for such a model (i.e. the mean plus spatial correlation) is always less than the value for the linear models without spatial covariance and is in some cases, significantly less. Some tables giving the GLM results for correlation function 1 with no transformation of the egg densities are presented here below.

Overall these results represent a classic example of the problem of fitting a spatial covariance model without accounting for possible trends and the problem of distinguishing between trends and covariance particularly with small data sets.

Table 1:

Maximum log-likelihood values from stepwise fitting of model 1 for correlation function 1 and for no transformation of the egg density estimates.

Parameters included	No correlation	Correlated errors
b_0	-131.44	-127.20
$b_0 + b_1$	-125.72	-125.72
$b_0 + b_2$	-126.42	-127.91
$b_0 + b_1 + b_2$	-125.72	-125.72

Table 2:

Parameter estimates from step wise fitting of mode 1 with correlation function 1.

Parameters included	Estimates of the parameters				
	b_0	b_1	b_2	L	σ^2
b_0	3.78	-	-	0.248	17.5
$b_0 + b_1$	0.42	3.61	-	0.00	11.8
$b_0 + b_2$	1.81	-	4.47	0.19	
$b_0 + b_1 + b_2$	0.42	3.60	0.014	0.00	14.7

b) Robert West: Fitting correlation function models and testing for the best fit, including a no correlation option.

Model

Given the small amount of data, the simplest of models is chosen: namely

density = $\mu + e$

where

μ *is an overall mean density*

and

e is an "error" term with mean 0, variance σ^2, *and is such that errors corresponding to two points separated by a distance d are correlated according to a function of d.*

It is also assumed that the distribution of egg densities is roughly normal. Since the data comprises large counts over an area, this assumption appears reasonable. An alternative would be to transform the data by taking the square root. This was not done so that this approach could more readily be compared to other methods.

Fitting

Maximum likelihood estimation was used to obtain estimates of parameters and compare models of the correlation function.

Results

Correlation Function	Log-likelihood
Uncorrelated	-131.436
Gaussian	-129.840
Hole effect	-128.887
Cauchy	-128.106
Spherical	-127.656
Exponential	-127.197

The log-likelihood is thus maximized by the exponential model: that is amongst the models considered the exponential model gives the best fit to the data. The likelihood-ratio test shows that all correlation models except the Gaussian are significantly better than the uncorrelated model at the 5% level. In particular the exponential model:

$\mu = 3.776$ (*1000 eggs per sq m)
$\sigma^2 = 17.66$ (*1000 eggs per sq m)^2

and

correlation at distance d (km) = $e^{(-d/\alpha)}$
$\alpha = 0.250$ km

is a significantly better fit than the uncorrelated model (for which $\mu = 3.802$ *and* $\sigma^2 = 17.45$, *the maximum likelihood estimate of being 28/29 of the usual (unbiased) variance estimate).*

Confidence intervals

from the information matrix of the maximum likelihood fitting procedure, we obtain a symmetric 95% confidence interval for the mean parameter μ

*(2.240,5.312) *1000 eggs per sq m.*

NB Since there is spatial correlation, the interval is larger than that for the mean if it could have been shown that the observations were uncorrelated (roughly twice in this case). That is if an analysis ignores the spatial correlation then the mean would be given with "false" accuracy.

From the examination of the profile likelihood for we get the following asymmetric 95% confidence interval for :

(0.102,1.080) km.

Nugget effect

A slightly better fit can be obtained by including a nugget effect in the model. However the likelihood ratio test shows this to be Not Significant.

Mapping

A map of the "patch" could soon be produced using one of the readily available Kriging software packages which allows the specification of the above correlation function. Note that the semivariogram equivalent is given by:

$$\Gamma(d) = 17.66\ (1.0 - \exp(-d/0.250)$$

To the author's knowledge, maps of the errors associated with the kriged estimates produced by these packages do not usually allow for the fact that the variogram has been estimated, thus to produce an acceptable error map, the author would suggest the generation of a series of simulations conditional upon agreement with the data values. This has not been done due to shortage of time before the submission of the present report.

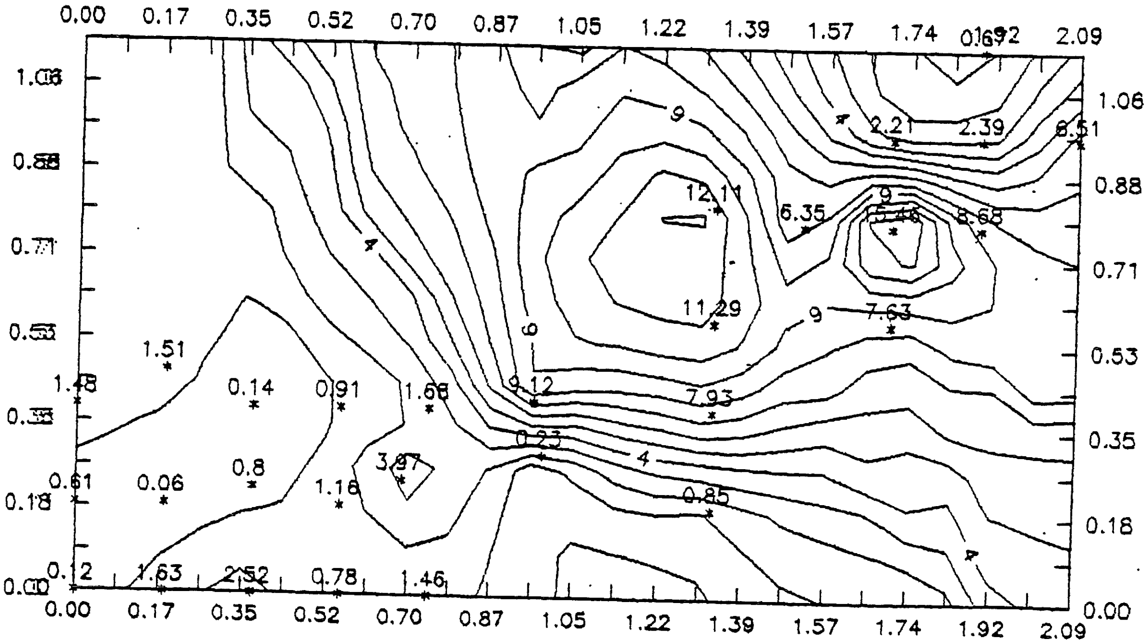
Total abundance

There is little problem in obtaining an estimate of total abundance condition upon the data complete with confidence interval . However this would involve the (possibly numerical) integration over the relevant area which was not given and time excluded this option. A more constructive approach would be to obtain this from conditional simulations as mentioned above. These would provide the entire distribution of the total abundance estimate and well as demonstrating and well as demonstrating the random process model.

Conclusion

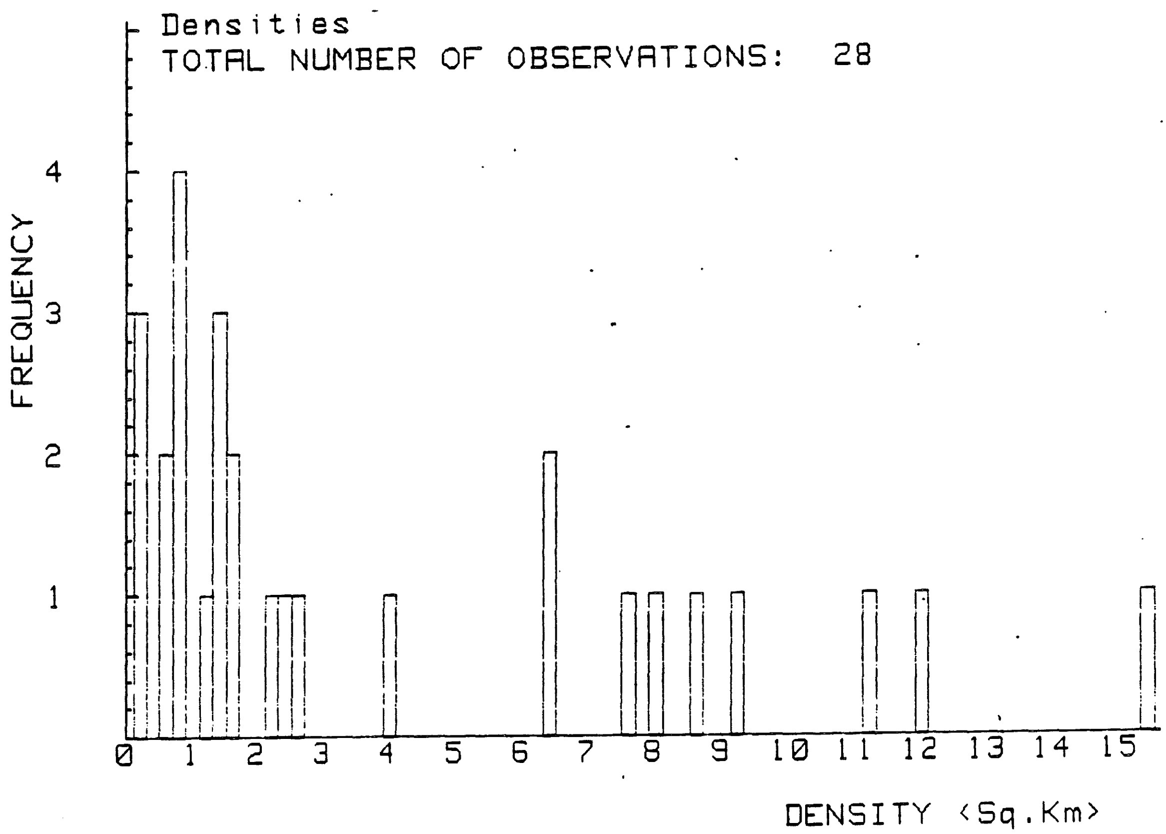
For our simple model the preferred fit to the data included a term for spatial correlation. That is from the limited data we found evidence of "spatial structure" which means that the data can NOT be regarded as being uncorrelated.

c) Yvan Simard provided a contour map generated by the surfer software based on kriging, but with a fixed variogram model:

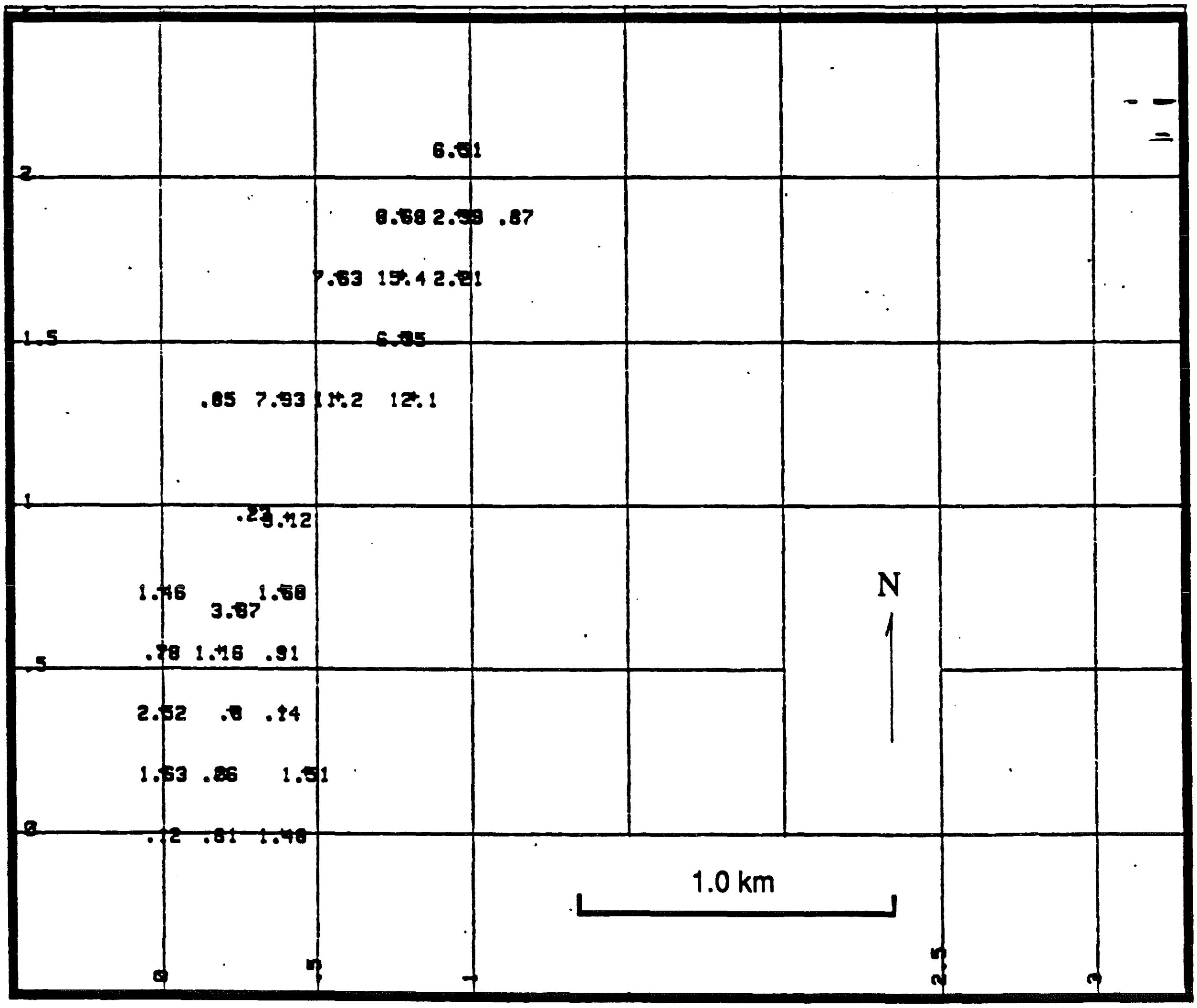


d) G.Y. Conan and Elmer Wade provided results from GULFKRIG software:

(1) Spatial dispersion of the sampling units

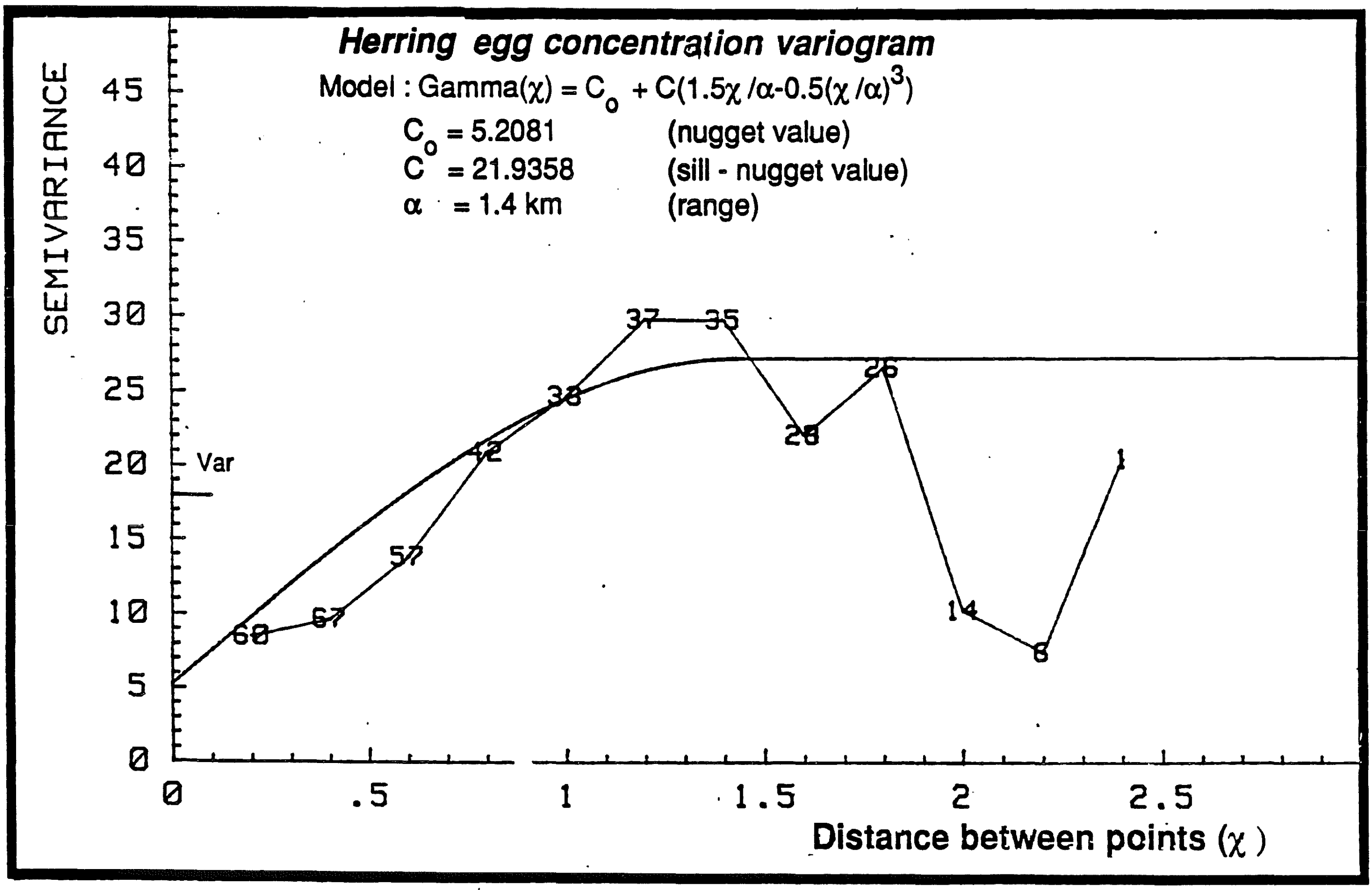


(2) Sampling distribution



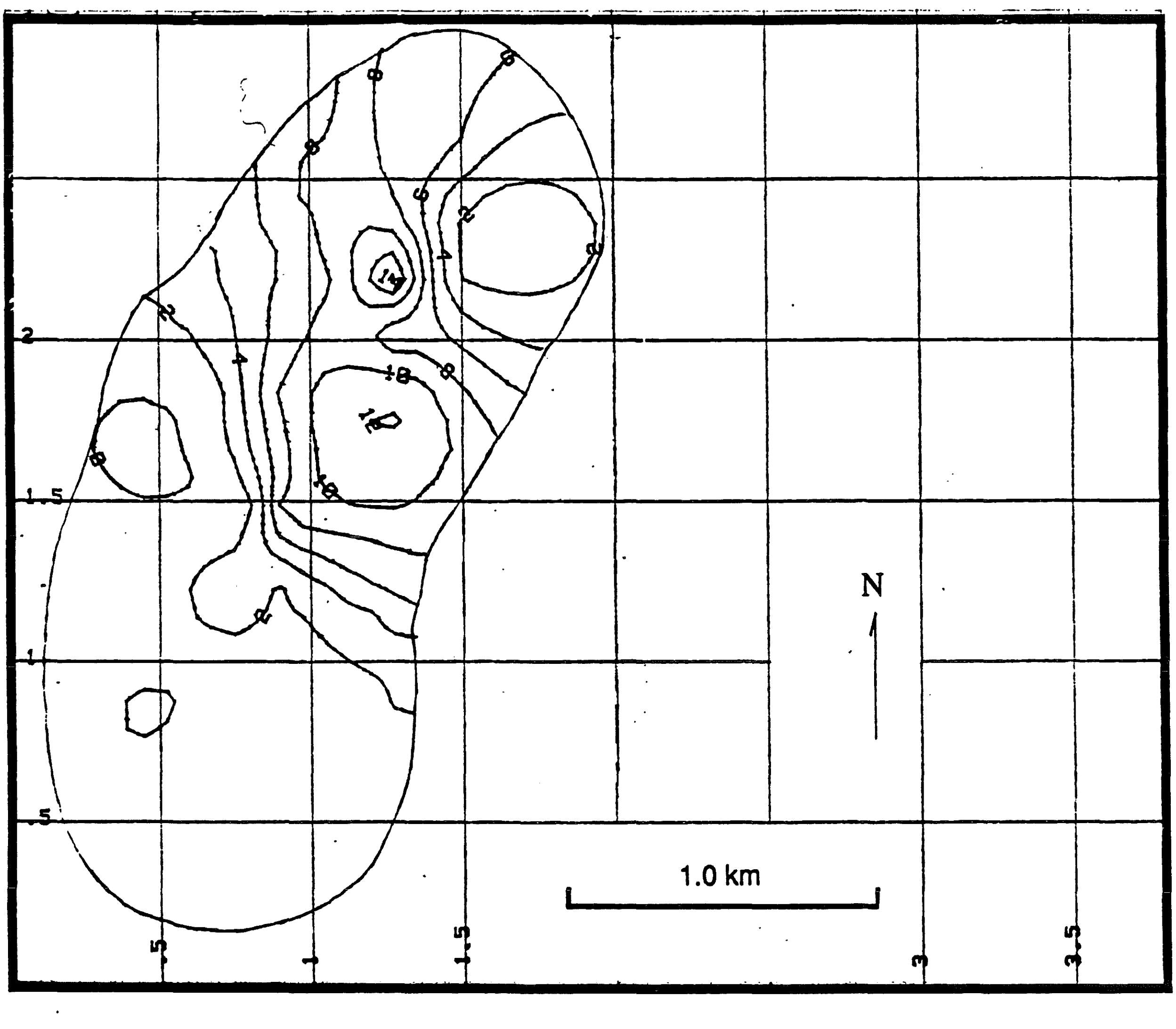
(3) Variogram.

As in most packages no attempt is made to fit a model to the experimental variogram by an automated statistical method such as maximum likelihood or least squares. The program allows to choose among several options: exponential, spherical, fractal, nested spherical. The fit is obtained by trial and error after digitizing on the screen reference points such as range of covariance effects, sill, and nugget effect. The drop in variance beyond the range is an effect frequently observed around a "patch" it occurs when the semi variance is calculated between low density points at the periphery of the domain, especially between zero's. The numbers on the variogram refer to the number of pairs encountered at the corresponding distance apart.

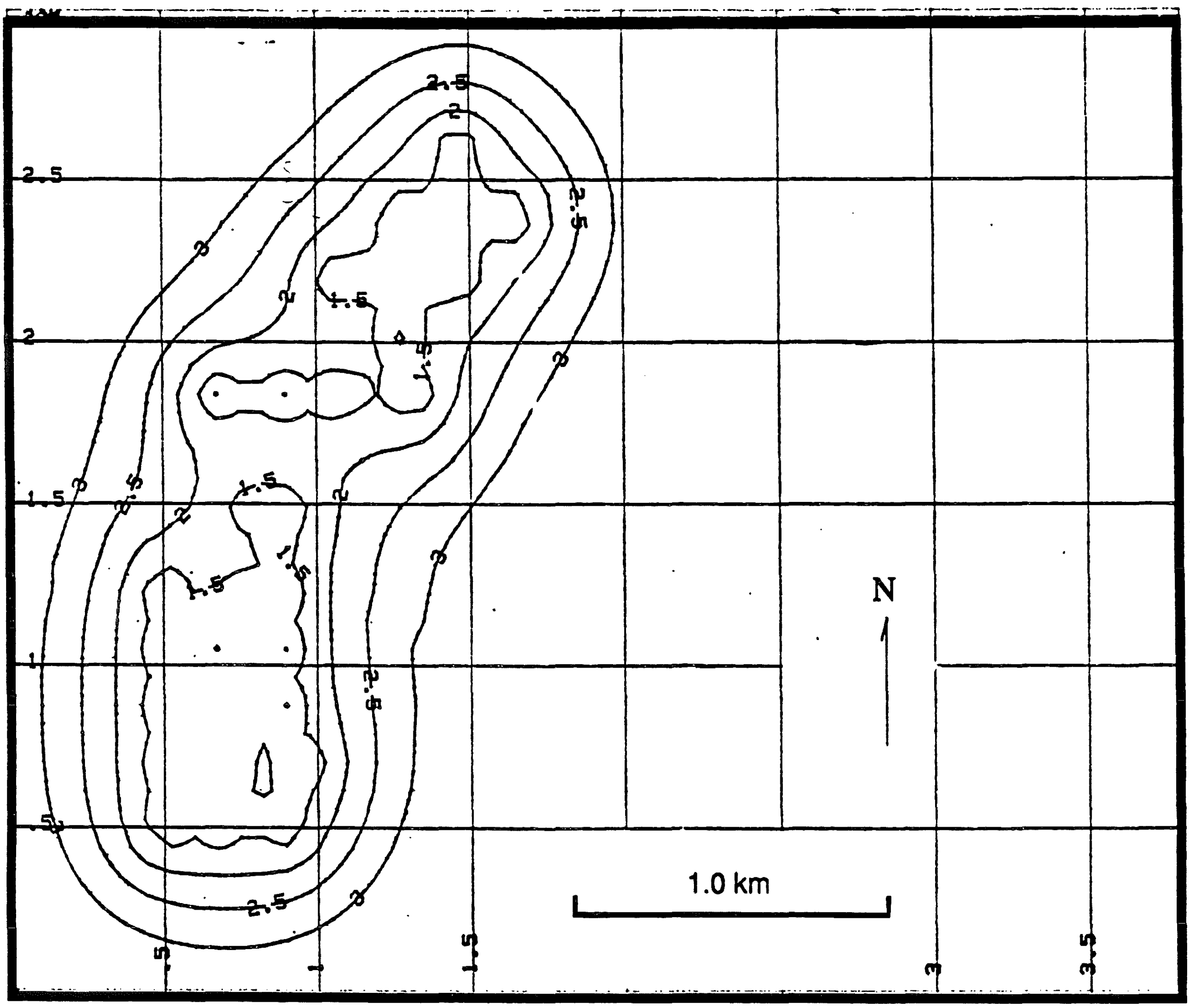


(4) Isodensity contours

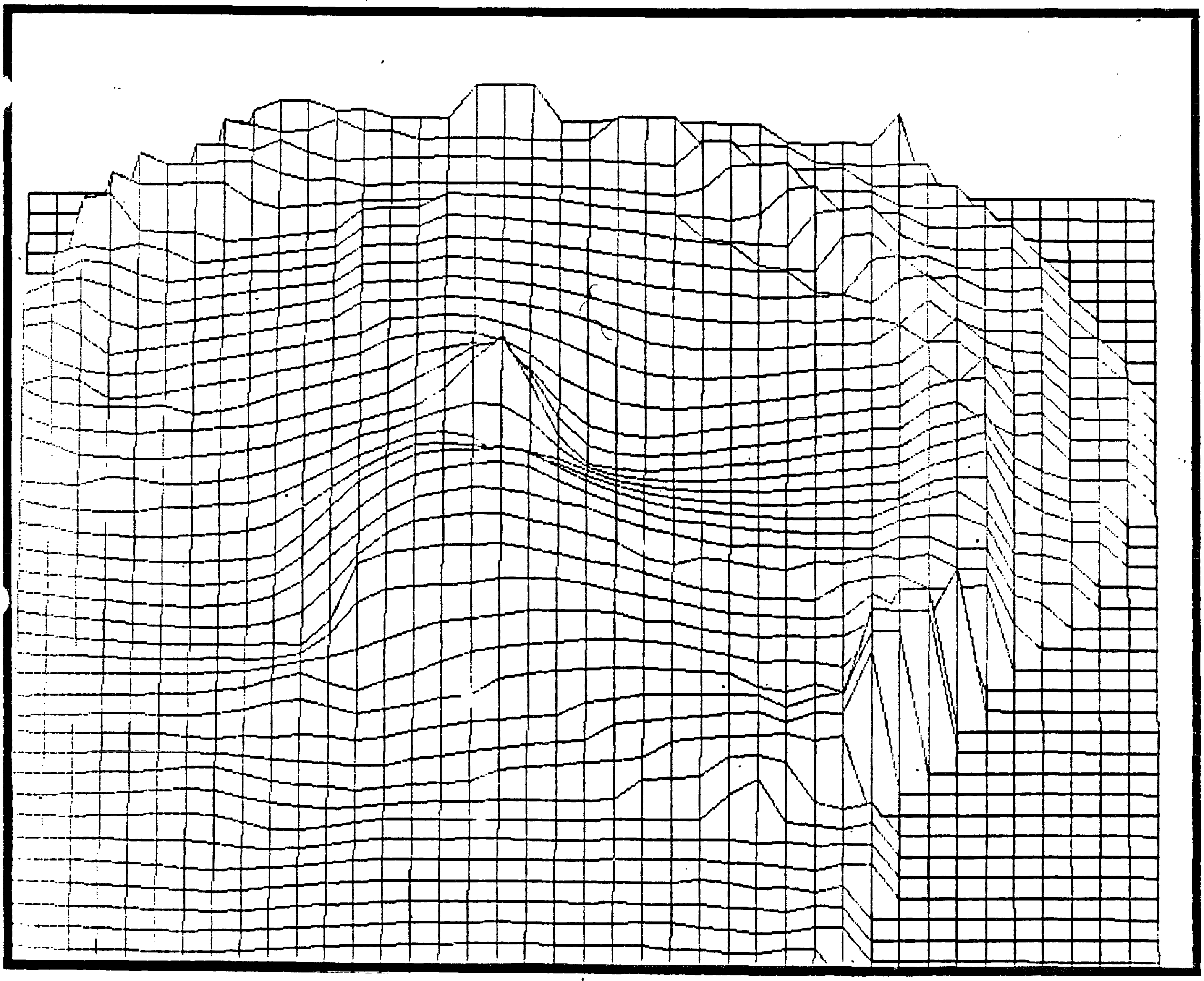
The contours are drawn only within the boundaries of a variance contour calculated from the kriging information in order to avoid extrapolating beyond reasonable limits.

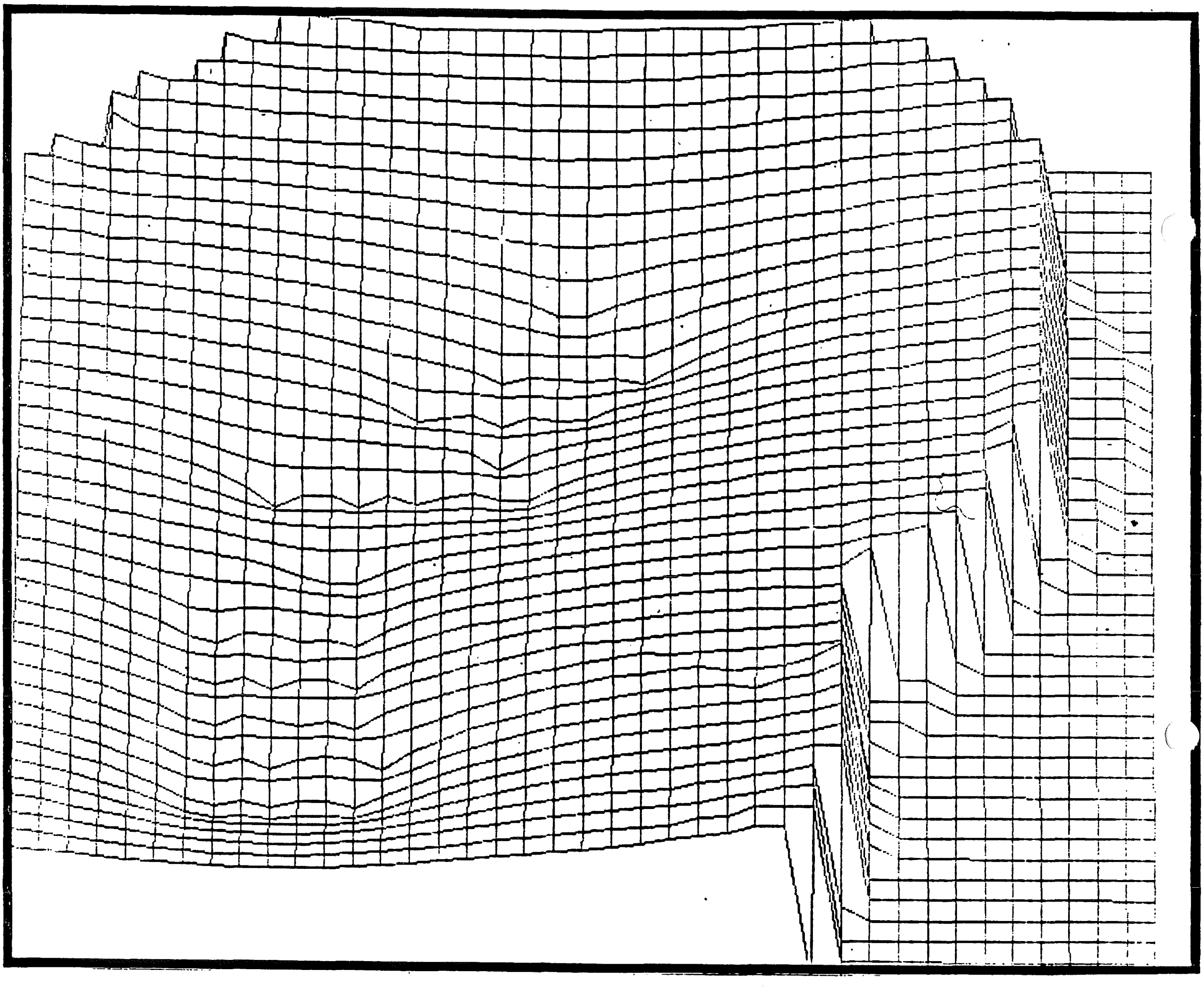


(5) Isovariance contours. These contours represent the estimation variance provided by kriging. The distance between each sampling unit and the distance from the unknown point to each of the sampling units are taken into account for calculating this variance. However the variance of estimation of the parameters of the variogram is not taken into account in these calculations.

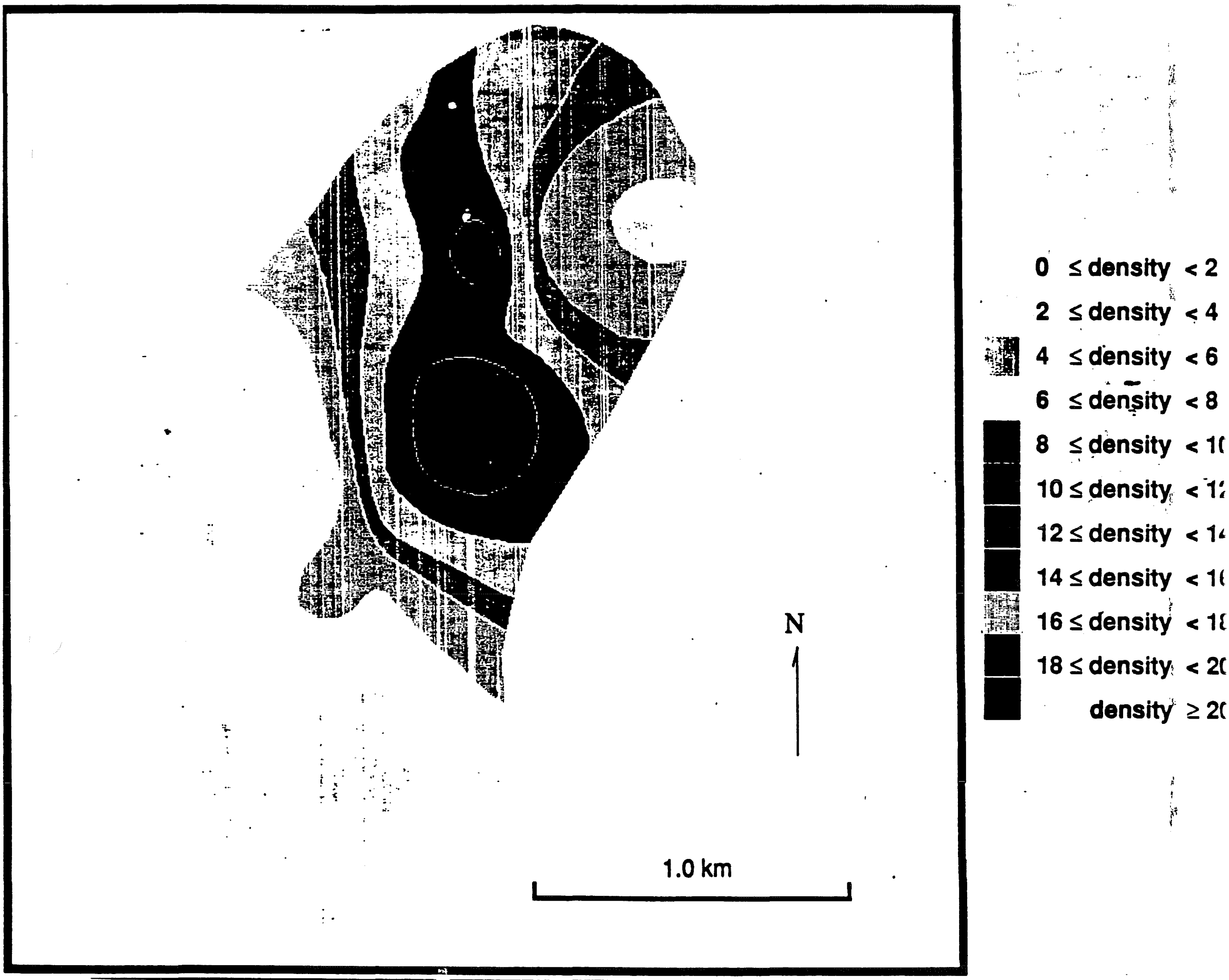


(6) Three dimensional representations of the density and variance contours. Plane contours are difficult to interpret, a 3D representation allows to better visualize the patterns in the structure.

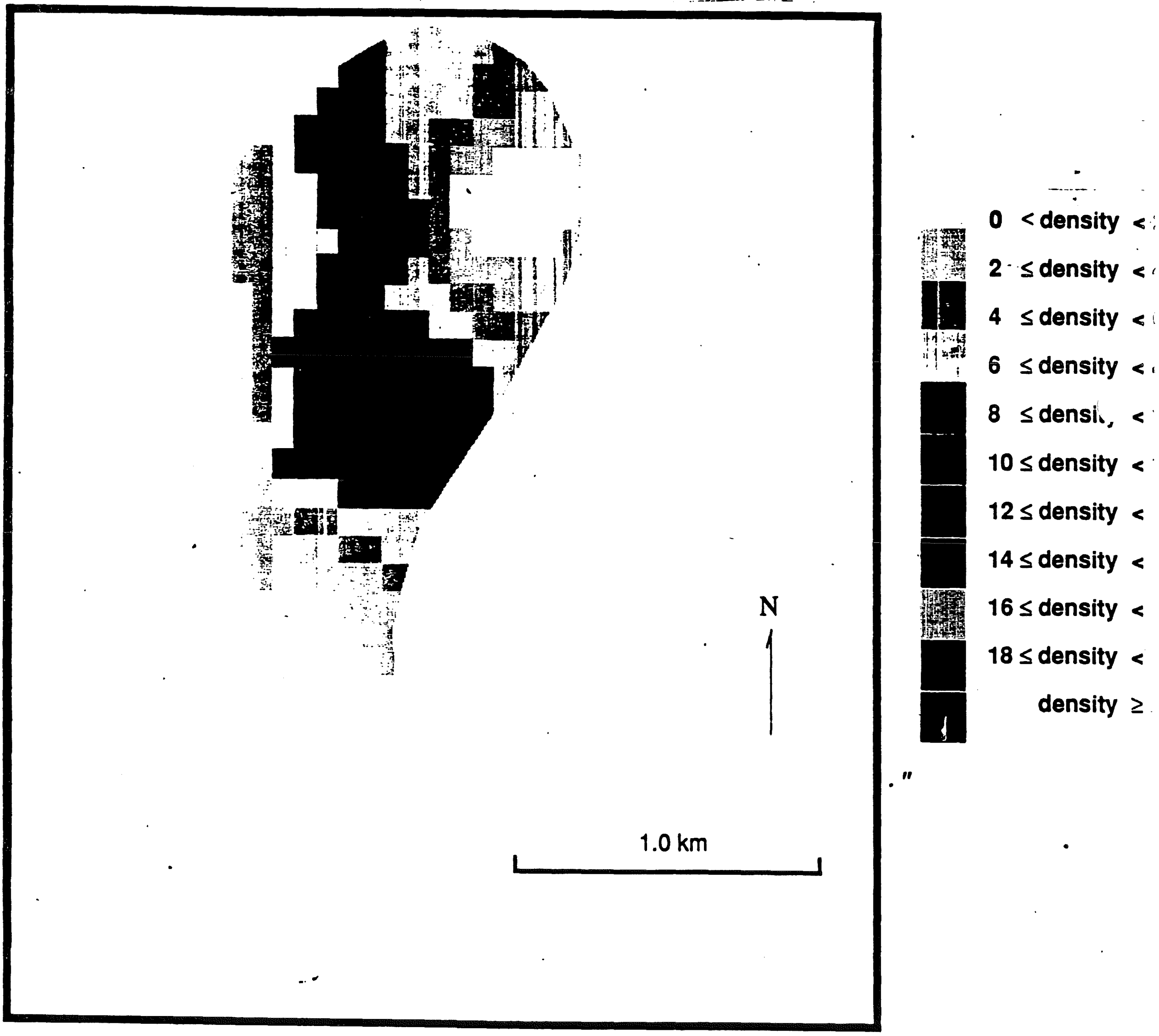




(7) Shaded or colour representations of the structure provide an easy quantitative interpretation at any required degree of resolution.



(8) *Block kriging representation of densities. An average estimate is provided within each square. This approach provides the possibility to estimate the portion of biomass which can be harvested above a cut off sill of local density set for biological or economic purposes. For instance it may be decided that harvesting of herring eggs for exportation is unsafe for maintaining recruitment in area where the density is lower than a cut off value of 10. The harvestable biomass may be estimated in contrast with the standing biomass by integrating over the whole area surveyed.*



(9) *Global density estimates. The concept of average density is meaningful only when it is associated with a a defined geographic domain. A problem generally overlooked is the geographic range of the resource. In this case study for instance the maps show that the sample points do not define the limits of the resource. For the purpose of this exercise the density was averaged within the contour of variance defining reasonably safe limits of information.*

Arithmetic mean:	*3.79*	*Variance of the mean:*	*.6197*
Area surveyed:	*3.234*		
Kriging average:	*3.85*	*Kriging variance:*	*.6519*

The area is in Km^2, the mean and average densities in 10^3 eggs/m^2

e) General remarks by G.Y. Conan

This case study illustrates the divergence of points of vue encountered during the meeting. One analysis emphasized the presence of a drift and advocated no covariance, three other analyses advocated the presence of covariance. The differences in interpretation originate mainly from the type of approximation model used. The estimates of density are actually quite similar.
By observation of the map of the resource provided by kriging, it can be easily understood why there was a divergence of opinion as to using a linear drift versus a covariance in the residuals. The patch is not exhaustively sampled, the sampling units towards the south actually missed the concentration of eggs creating a significant drift. Should the sampling locations have been distributed differently, the drift would not have been apparent. This illustrates the fact that a statistical test should always be associated with a visualization of the data. Within the range of observations data sets can be modelled satisfactorily according to different interpretations. However when it comes to extrapolation the results may widely differ between models.

Authors 2 and 4 agree that the variance of the arithmetic mean in the presence of covariance is likely to provide overoptimistic estimates of the precision. The approach presented by author 2 for fitting a model to the covariance data is statistically elegant, but could nicely be complemented by a scattergram of residuals in order to detect visually departures from the model towards the range limits of the data.

Authors 2 and 4 have perceived a difficulty in defining the limits of the geographic domain of the data. This problem may be easily overlooked during standard surveys if mapping is not conducted simultaneously in order to adapt the sampling design as a function of the information received. The observation of the density contours show that the present survey could be optimised if new samplin units were located towards the NW and SE of the existing ones rathe than towards the SW.

Authors 1 and 2 use a cautious approach and set great importance on testing the significance of models by reference to a null hypothesis. Authors 4 do not set a null hypothesis, but search for a reasonable fit of a reasonable model. They stress that there is no directing a priori null hypothesis a model can be tested against in this particular case study, they suggest that there is no logical reason to chose the absence of spatial autocorrelation as a reference null hypothesis. In the absence of external information, the best model is the one which best fits the data, whether or not the departure from an arbitrarily preset alternative is statistically significant.

6 DATA SETS

The following data sets were available for analysis during the meeting using software packages provided by the participants . However, to the request of several of the holders, these sets, many of which yet unpublished, were not placed in the public domain and the results were not listed or recorded.

Snow crab Trawl Survey, Gulf Region DFO Canada.

Pandalus borealis, IML, DFO, Canada.

Herring Acoustic Survey, Gulf Region DFO Canada.

Chlamys opercularis and *Nephrops* Survey, Dredge, Photography. Fiskirannsoknarstovan, Faroe Islands.

Multispecies Acoustic Survey, Lac d'Annecy, INRA, France.

Hake Trawl Survey, Bay of Biscay. IFREMER, France.

Nephrops and Young Fish Survey, Havsfiskelaboratoriet, Sweden.

Ostrea edulis, MAFF Lowestoft, U.K.

Multispecies Trawl Survey Data, NEFC NOAA USA.

Surf Clam, NEFC NOAA USA.

7 SOFTWARE DEMONSTRATIONS

The following sets of software were available for demonstration at the meeting for processing data sets provided by the participants.

BLUEPACK. Sun UNIX workstations. Kriging package based on Intrinsic Random Functions of order k theory. The software can cope with stationary variables as well as phenomena showing systematic trends, it can provide point estimates, average values over rectangular grid cells, average values over a territory of any shape. Developed by Centre de geostatistique de Fontainebleau, available from Geovariances 1 rue Charles Meunier 77210 Avon-Fontainebleau France.

GEO-EAS. IBM Compatible MS DOS workstations. Kriging package designed mainly for production of 2-dimensional grids and contour maps of interpolated (point kriging) estimates from sample data. Variogram computation and fitting to various models. Developed by US Environmental Protection Agency with the counselling of Andre Journel (Stanford University). Available at no cost from E. Englund USEPA, Environmental Monitoring Systems Laboratory, Las Vegas, Nevada 89119.

GLM Methods and IMSL routines. VAX, fortran 77. Variogram fitting by optimised methods, testing of significance of spatial covariance. Available from NMFS/NOAA, Woods Hole laboratory, USA.

GULFKRIG. HP UNIX or HP BASIC series 300/200 workstations. Standard stationary kriging package provides variogram fits, contour maps from point kriging, averaging over cells in a grid (block kriging) and within irregular contours. Converts LORANC and Geographic coordinates to square grid coordinates. Designed for marine resources estimates. Adapted by Gerard Y. Conan and Elmer Wade from a mining engineering package available from Lynx corporation, Vancouver, Canada.

MAGIK. IBM compatible MS DOS workstations. Intrinsic random function package designed for mapping and contouring. Particularly adapted for teledetection surveys from satellites or aircrafts. Designed by Francis Gohin and J.J. Lechauve, IFREMER and ORSTOM, France. Available from Geovariances, 1 rue Charles Meunier, 77210 Avon-Fontainebleau, France.

R Package. Apple Macintosh workstations, Pascal. Multivariate and spatial analysis. Clustering methods, contingency periodogram, chronological clustering, spatial autocorrelation tests, Mantel tests, COCOPAN. Developed by Alain Vaudor and Pierre Legendre, Departement de Sciences Biologiques, Universite de Montreal, CP 6128, Succursale A, Montreal, Quebec, Canada H3C 3J7. Available at no cost.

SSDSS SPLINE SURVEY DESIGNER SOFTWARE SYSTEM. IBM compatible MS DOS workstations. Software designed on the basis of spline approximation methodology for mapping a resource, estimating its abundance and optimising sampling during a survey. Incorporates depth information as an interpolating factor. Designed by Dimitri Stolyarenko, VNIRO, 17 Verkhne Krasnoselskaya, VNIRO, Moscow, USSR.

SURFER. IBM compatible MS DOS workstations. Includes a contouring routine based on point kriging using a fixed model variogram. Available from Golden Software, 807 14th street PO Box 281, Golden, Colorado 80402, USA.

8 PAPERS PRESENTED

Conan, G.Y., U. Buerkle, E. Wade, M. Chadwick, and M. Comeau, 1988. Geostatistical analysis of spatial distribution in a school of Herring. Unpublished manuscript.

Conan, G.Y., M. Moriyasu, E.Wade, and M. Comeau, 1988. assessment and spatial distribution surveys of snow crab stocks by geostatistics. Unpublished manuscript.

Legendre P. and M.J. Fortin, 1989. Spatial pattern and ecological analysis. in press in Vegetatio.

Legendre P., N.L. Oden, R.R. Sokal, A. Vaudor, and J. Kim. 1989. Approximate analysis of variance of spatially autocorrelated regional data. Submitted to the Journal of Classification.

9 BIBLIOGRAPHY

Clark I., 1979. Practical Geostatistics. Elsevier London, 129 pp.

Cliff A.D. & J.K. Ord, 1973. Spatial autocorrelation. Pion Ltd. 178 pp.

Conan G.Y., D. Parsons and E.Wade. 1989. Geostatistical analysis, mapping and global estimation of harvestable resources in a fishery of Northern Shrimp (*Pandalus borealis*).

Elliot J.M., 1971. Some methods for the statistical analysis of samples of benthic invertebrates. Scient. Publs. Freshwat. Biol. Ass. U.K. 25: 144pp.

Guillard, J.D. Gerdaux, and G.M.Chautru, 1987. The use of geostatistics for abundance estimation by echointegration in lakes: the example of lake Annecy. International Symposium on Fisheries acoustics June 22-26, Seattle, Washington, USA.

Hardy P.L. , 1984. Universal kriging, least squares collocation, and optimum interpolation (what's the difference?) manuscript presented at the 16th European Meeting of Statisticians (EMS), University of Marburg, Germany 3-7 September 1984. 10 pp.

Journel, A.G., & Ch.J. Huijbregts, 1978. Mining Geostatistics. Academic Press, 600p.

Matheron G., 1971. The theory of regionalized variables and its applications. Les cahiers du CMM, fasc. 5. ENSMP, 211 p.

Messieh S., 1988. Spawning of Atlantic Herring in the Gulf of St. Lawrence. American Fisheries Soc. Symposium 5:31-48.

Messieh S.N., and H. Rosenthal, 1989. Mass mortality of herring eggs on spawning beds on and near Fisherman's Bank, Gulf of St. Lawrence, Canada. Aquat. Living Resour., 2,1-8.

Peitgen H.O. & D.S. Saupe eds. 1988. The Science of Fractal Images. Springer-Verlag New York. 312 pp.

Stolyarenko, D.A. 1987. The spline approximation method and survey design using interpolation with a microcomputer: Spline Survey Designer Software System. ICES, Doc. C.M. 1987/K:29.

Stolyarenko, D.A. 1988. Local integral measurements: an advanced spline approximation method for trawl, echo sounding and television surveys. ICES, Doc. C.M. 1988/D:1.

Stolyarenko, D.A. 1989. A geographic information system for fish surveys. ICES, Doc. C.M. 1989/D:3.

10 CONCLUSIONS

The multidisciplinarity of the workshop was well received by the participants and generated challenging discussions which were at times difficult to follow by all participants due to the lack of standardized vocabulary in spatial statistics.

No consensus was met on the opportunity to substitute regular grid sampling to random or stratified random sampling for the global assessment of a resource. However most participants were appealed by the principle of correcting a biased sampling design by using spatial approaches such as kriging , splines or polynomial approximations. It was suggested that this may be an approach for using commercial fishery data for assessing a resource.

No consensus was met on the application of fractal models or even on their definition for assessing a resource.

The concept of incorporating parameters other than geographic coordinates for establishing the statistical distance between two points was considered with interest for future research.

The step of fitting an optimised model to the experimental variogram in the kriging methodology was considered critical by some but of limited importance by others.

The presence of drifts in data sets was considered as an obstacle to fitting a variogram in the present methodology by some participants who strongly advocated further research in this field, while others felt comfortable with the existing methodology.

The ability of developing regionalized (local) estimates for a resource, allowing for economic considerations as to the benefits of fishing in a specific location was considered with interest.

The concept of intrinsic random functions of order k was considered as mathematically elegant by many participants, but its complexity compared to standard kriging was questioned by others.

For most participants the workshop was more an initiation to unfamiliar concepts, techniques, or data than an opportunity to process real life data for assessment needs. It was generally felt that the workshop had been a challenging forum for new and promising ideas, and that it should be reconducted in order to allow for the processing of full data sets, and to develop new methodologies and concepts based on worked out case studies in assessing and mapping fisheries resources.

11 RECOMMENDATIONS

1) Survey design

The problem of designing surveys was considered crucial. Much of standard statistics is invalidated in the case of spatial correlation. Spatial statistics approaches for design of surveys involving mapping and assessing were discussed and considered promising after the presentation of methodologies such as Spline Survey Designer Software System (VNIRO USSR). Transition to 3 or more dimensional models involving non stationary probability curves is of great interest and should be given further consideration.

2) Acoustic surveys

Acoustic surveys of squid and shrimps are currently being developed.The methodology of acoustic surveying of fisheries was generally considered.

The consensus is that, when preliminary surveying is not feasible, or when other particular prior knowledge about the stock's distribution is not available, the acoustic sampling could be satisfactorily done along parallel, equally spaced line transects. In general, the transects should be crossing with the maximum density gradient. Resulting line measurements of resource density should be interpolated to form an area density distribution according to kriging principles, or other spatial statistical techniques such as splines, least square trend surfaces, inverse square distances. Some of these techniques allow immediate computation of an unbiased estimate of the mean with associated estimates of variance and confidence limits.

3) Resource patchiness and kriging

The frequent extreme inhomogeneity of marine resource occurrence, as due to clumping and schooling, is recognized to require special care in the application of kriging. It is suggested that the Shellfish Committee call the attention of the Statistics Committee on this matter and recommend future joint meetings and appropriate working group fora.

4) Reconduction of mandate

This informal group recommends that the workshop be turned into a formal working group, that should meet again to discuss further developments in spatial statistics and their use in fisheries research.

12 ACKNOWLEDGMENTS

The participants and the chairpersons wish to thank the French Delegate Alain Maucorps for proposing to host the workshop on the Campus of Centre Oceanologique de Bretagne, IFREMER Brest, as well as the director of the centre Dr Marthe Melguen, the chief of the Computer department M Francois Leverge, Miss Brigitte Millet and MM

Patrick Berthou and Francis Gohin for their most competent help and their kind hospitality.