

Northwest Atlantic



Fisheries Organization

Serial No. N4899

NAFO SCS Doc. 03/22

SCIENTIFIC COUNCIL MEETING - SEPTEMBER 2003

**Report of the Workshop on Mapping and Geostatistical Methods
For Fisheries Stock Assessment**

by

D. W. Kulka¹, L. Hendrickson², N. Bez³, R. Schlitzer⁴, G. A. P. Black⁵, and M. R. Simpson¹

¹ Science, Oceans & Envir. Br., Dept. of Fish. & Oceans, 166 Water St.,
P.O. Box 5667, St. John's, NL A1C 5X1, Canada

² National Marine Fisheries Service, NEFSC, Woods Hole, MA 02543, USA

³ Centre de Géostatistique, Ecole des Mines de Paris, 35 rue Saint Honoré, 77305 Fontainebleau, France

⁴ Alfred Wegener Institute for Polar & Marine Research, Columbusstrasse, 27568 Bremerhaven, Germany

⁵ Invert. Fish. Div., Mar. Fish Div., Fisheries & Oceans, P. O. Box 1006, Dartmouth, NS B2Y 4A2

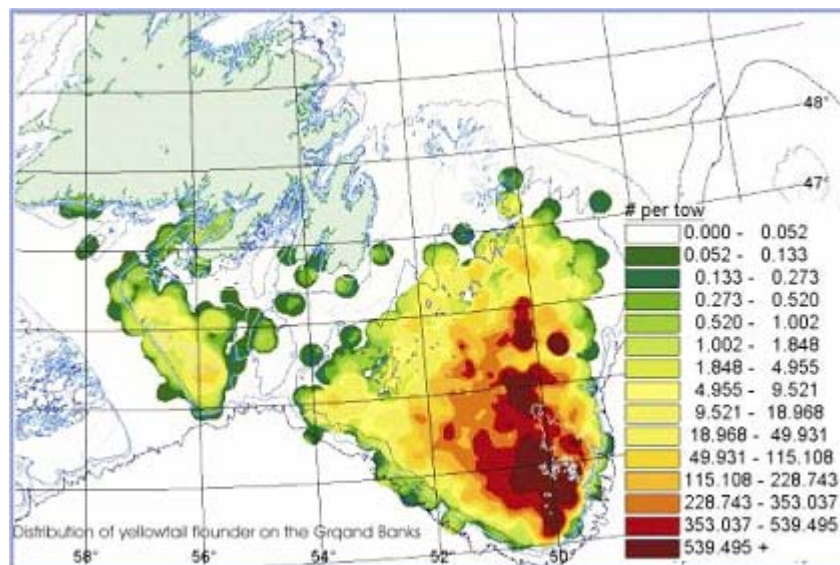


TABLE OF CONTENTS

1.0	INTRODUCTION	3
2.0	WORKSHOP OBJECTIVES AND OVERVIEW	3
3.0	SPATIAL VISUALIZATOIN AND ANALYSIS	3
3.1	Mapping Software and Techniques	4
3.1.1	ACON	4
3.1.2	SPANS	9
3.1.3	Ocean Data View	12
3.1.3.1	Operational Modes	12
3.1.3.2	Gridding Algorithms	13
3.1.4	Generalized Additive Models	14
3.2	Demonstration - Solving a Spatial Question	15
3.2.1	Ocean Data View	17
3.2.2	R-Geostatistics	18
3.2.3	ACON	18
3.2.4	SPANS	19
	UTILIZATION OF GEOSTATISTICS FOR MAPPING	20
4.1	History, theory and concepts	20
4.2	Consideration of the spatial aspects of fish distributions	21
4.3	Identi fication of spatial structure using variograms and kriging	21
4.4	Methods for summarizing spatial distributions	22
4.5	Other geostatistical methods	22
4.6	Geostatistical analyses of Grand Banks data sets	22
4.6.1	Yellowtail flounder	22
4.6.2	Depth	25
4.6.3	Greenl and halibut	26
4.6.4	CPUE	27
5.0	CONCLUSIONS	28
6.0	REFERENCES	28
7.0	ACKNOWLEDGEMENTS	28
APPENDIX 1.	WORKSHOP AGENDA	30
APPENDIX 2.	WORKSHOP PARTICIPANTS	33

1.0 INTRODUCTION

Geographic visualization and geo-spatial analysis of fish, fishery and environmental data are increasingly more important components of fisheries research. Papers incorporating these techniques are becoming more common in the work of Scientific Council, this proliferation aided by the availability of GIS software. Such application tools provide mapping functions, and in addition, many also have modelling and geostatistical functionality.

In marine environments, fish and invertebrates are not distributed at random but are organized in space and fisheries and research data are often associated with a geo-reference, usually latitude and longitude. This organization results in the realization of a variable such as fish abundance being proximally related, i.e., geographically referenced data are spatially correlated. Classical (non-spatial) analyses used to compute confidence limits of a variable assume that error terms of samples are stochastically independent of one another. This condition is often not met with spatial data. Geostatistics were developed to deal with estimation problems in spatially-correlated data.

To address the recommendation of Scientific Council a Special Session, “Workshop on Mapping and Geostatistical Methods for Fisheries Stock Assessment”, was held at the Holiday Inn in Dartmouth, Nova Scotia during 10-12 September 2003 convened by David Kulka (Canada) and Lisa Hendrickson (USA). Instructors included: Dr. Nicolas Bez (Centre de Géostatistique, Fontainebleau, France), Dr. Reiner Schlitzer (Alfred Wegener Institute for Polar and Marine Research, Bremerhaven, Germany), and Jerry Black, Dr. Mark Simpson, and David Kulka (Department of Fisheries and Oceans, Canada). There were 28 participants from Canada, Denmark (Faroe Islands and Greenland), Estonia, the European Union (France, Germany, Portugal, and Spain), the Russian Federation, and the United States of America. Ralph Mayo (USA), Chair of Scientific Council, opened the workshop by welcoming the participants and thanking the co-conveners.

2.0 WORKSHOP OBJECTIVES AND OVERVIEW

L. Hendrickson presented an overview of the agenda and objectives of the workshop, including an introduction of the instructors and their technical backgrounds. The objective of the workshop was to provide participants with a basic understanding of geostatistical concepts and methods, as well as the tools to apply this knowledge to fisheries stock assessment through the use of hands-on exercises that incorporate, using freeware mapping and geostatistical software. Two data sets centered on the Grand Banks, Canadian spring bottom trawl survey sets from 1996-2002 (hereafter referred to as the survey data) and information from the yellowtail flounder fishery (hereafter referred to as the commercial fisheries data), were used for the lessons and demonstrations. The Canadian bottom trawl surveys constitute a stratified random design.

The first day of the workshop focused on the visualization and analysis of spatial data using GIS (Geographic Information Systems) software packages. Participants viewed demonstrations of the software and completed hands-on exercises. Geostatistical concepts and methods, particularly related to kriging, were presented during the following two days. Participants learned the basics of using the “R” software package (freeware) to conduct geostatistical analyses using routines prepared by programmers at the Centre de Géostatistique as well as those developed during the workshop. The agenda, relevant literature citations and links to the software utilized in the workshop can be found in App. 1.

3.0 SPATIAL VISUALIZATION AND ANALYSIS

D. Kulka highlighted the long standing need for geographic referencing of information and spatial visualization and analysis of fisheries resources and fisheries data. For example, the 1921 Proceedings of NACFI (North American Council on Fishery Investigations), a precursor of NAFO, wrote about “the need to represent the nature of grounds, to make evident at first sight the distinctive characters of the region in question from a sedimentary viewpoint, because in fishery matters the importance of sediments cannot be denied and to represent by conventional signs the bottom fauna; determine the location of whelk and shell fish beds that are sometimes utilized for the production of bait...”. They further indicated that “maps which have been published are appreciated by the fishermen, to whom they bring much information”. Geographical representation, then and now, is perhaps even more important to fishery scientists.

Background on the spatial structure of data was presented: geographic features are represented as point, vector or raster (described as area, polygon or surface) data. Points on a map represent locations for geographic entities such as sampling sites and can be used to spatially present measured variables such as average temperature or catch rate. Usually in the case of fishery analyses, the raw data comprise points (fishing or sampling sets) with latitude and longitude as the geo-reference plus various attributes describing various biological and physical components. Vector data consist of a set of connected points (lines and arcs are synonymous when used in this context), normally used to represent physical geographic features. Raster data are used to represent a region enclosed within a boundary. An area or series of related areas can be used to represent such features as a stock distribution or a temperature surface. These various geographic components can be overlain in various combinations in a map composition then analysed to determine the relationships among layers.

3.1 Mapping Software and Techniques

Data visualization and analysis was demonstrated using various types of GIS/geostatistical software including R (freeware), Ocean Data View (freeware), and SPANS. Attributes such as a catch rate or environmental variables in space (referred to as Z values) may be represented in their simplest state by classified points (such as expanding symbols) but for the purpose of analyses, surfaces (rasters covering the entire area of interest) provide added value through enhanced visualization and by facilitating further analysis of spatial relationships in the data. Several surface generating techniques were described and demonstrated using various software packages.

3.1.1 ACON

ACON (A CONtouring package) freeware (Black, 2002), an application developed to transform and visualize survey and commercial fisheries data was demonstrated by J. Black. The software can generate two and three-dimensional graphics and contains routines for generating maps in a number of projections. An internal scripting language (with an embedded compiler based on Extalk, Betz, 1988) is used to generate and execute virtual machine code. The language supports vector and matrix math in up to three dimensions.

Input data may be read from ASCII files, ODBC (Open Database Connectivity) data sources, or from Oracle SQL (Structured Query Language) databases. Output may be generated in graphics windows, as output files (e.g. PNG, Illustrator, PDF, PostScript, JPEG, SVG), or as movies in AVI or QuickTime format. The world vector coastline is provided as a Regionally Accessible Nested Global Shoreline structure at five levels of resolution (e.g. 0.1, 0.2, 1.0, 5.0 and 25 km). Bathymetry polygons are provided for the east coast of Canada, as are the NAFO convention area boundaries.

The transformation and visualization of point patterns to surface distributions was demonstrated using ACON through the Voronoi polygons and Delaunay triangulation tessellation methods. Voronoi polygons are generated from the circumcenters of natural neighbour circles formed between adjacent sampling points. Each Voronoi polygon edge is a perpendicular bisector of the matching Delaunay triangle (Fig. 1).

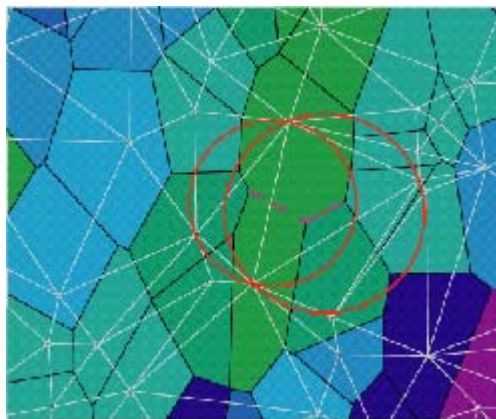


Fig 1. Delaunay Triangles (grey) and the matching Voronoi Polygons (black) with 2 circumscribing circles.

Delaunay triangles sub-divide the data region into triangles that are as equilateral as possible. Voronoi polygons are unique, as are Delaunay triangles except in the case of regularly gridded data. The relationship between Delaunay triangles and Voronoi polygons for the surveys was demonstrated (Fig. 2).

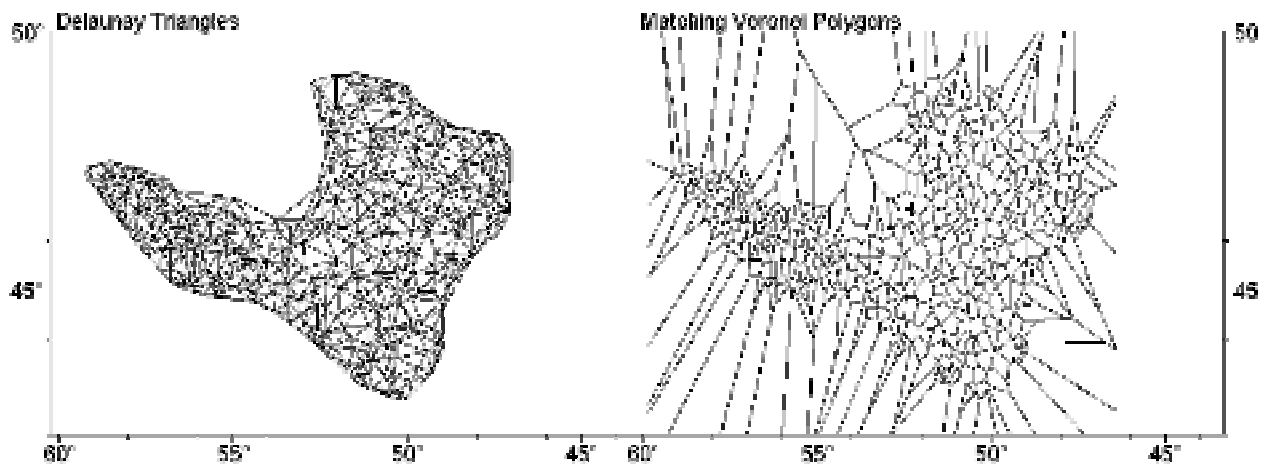


Fig. 2. Relationship between Delaunay triangles and Voronoi polygons for the research surveys. Left panel shows the Delaunay triangles formed from the 519 sets in the 1996 survey data, and the right panel shows the matching Voronoi polygons.

The pros and cons of interpolation methods for shaded surface contour generation such as moving averages, inverse distance weighting, and Voronoi analysis were presented. For example, the moving average method does not account for either the distance between neighbouring samples or between the samples and the target. In other words, all samples in a particular neighbourhood receive the same weight. With respect to a Voronoi analysis, all of the weight assigned to a particular data point is placed on its single-most, proximate neighbour. As a result, the spatial arrangement of the resulting Voronoi diagram is more a reflection of the location of the sampling stations rather than the values of the data at those stations. In addition, the limits of border samples may be questionable. The inverse distance weighting method assumes that the weighting factor is a function of the distances between data points and the target within a particular neighbourhood.

As an example of how custom output can be produced, a script was provided that generated a number of presentation graphics. The script demonstrated scaled symbol output of the number of Atlantic cod caught per std. (standard) tow (0.8 n. mi.) from the survey and surface contouring using inverse distance weighted gradient interpolation of the bottom temperature measured at each station (Fig. 3). Bathymetry contours were rendered as shaded contours and then overlaid with the observed bottom temperature as a shaded surface. Bathymetry contour lines were drawn over each surface to allow the user to infer the bottom topography in the region where the bathymetry contours were obscured by the bottom temperature surface. The inverse distance weighted gradient surface interpolation provides a visually similar shaded surface to potential mapping, in this case where the sampling density is high and a corresponding small scale map is used. In this example both the bottom temperature observations and Cod catch were aggregated to simple mean values per 15 minute square of longitude, latitude. The shaded coastline which overlays the bottom temperature surface was generated from the world vector shoreline stored in the RANGS file format. The Cod catches were plotted using a constant log ratio scaling (Bertin 1981) and rendered so that secant symbols remain distinct through ordered rendering white “halos” surrounding each symbol.

The script also demonstrated scaled symbol output of the catch of yellowtail flounder per std. tow using a similar graphic style (Fig. 4). This graphic was demonstrated through interactive browsing of the data set with *Data Dialog* windows (Fig. 4). Using this technique the data set was browsed interactively, by selection of the survey year, species of interest, and sampling metric (number or weight caught). The map included a Variogram as an annotation, and was automatically calculated for the selected catch.

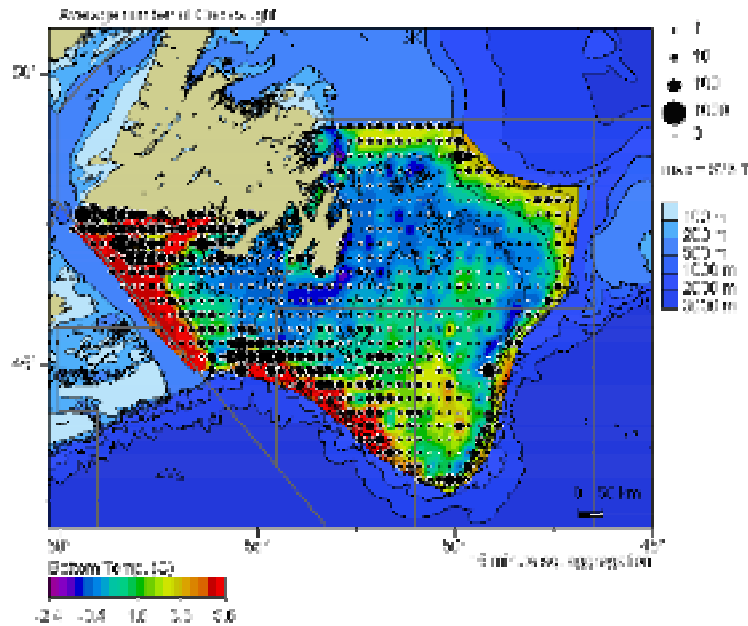


Fig. 3. Scaled symbol output (black dots) showing the distribution of Atlantic cod (number per tow) from 1996 to 2002 overlain on a surface contour of bottom temperature created using inverse distance weighted gradient interpolation of bottom temperature measurements.

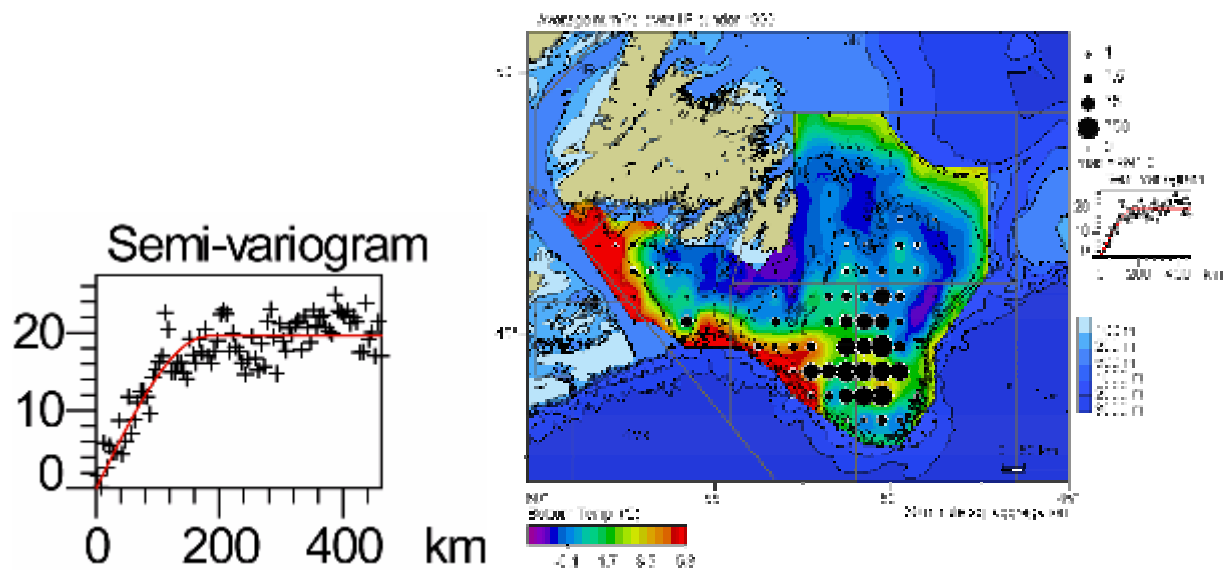


Fig. 4. Enlarged Variogram (of the plot located to right of map), generated through the *Data Dialog* window for the bottom temperature surface, for 1996 survey data. Catches of yellowtail flounder (number per tow) are overlaid as scaled symbols.

The plotting of multiple species on a single map using scaled symbol pie charts was demonstrated for the survey data to determine locations of cod and yellowtail flounder bycatch in the American plaice fishery (Fig. 5). This technique demonstrated one approach to illustrating the co-occurrence of species, and generates a synoptic view of the overall (and generally disparate) distribution, species specific abundance and combined abundance. Alternative levels of precision in the choice of data aggregation level generate variable levels of detail in the spatial distribution

plot that provide a trade off between a synoptic view and more precise detail in location information at the expense of increased information complexity.

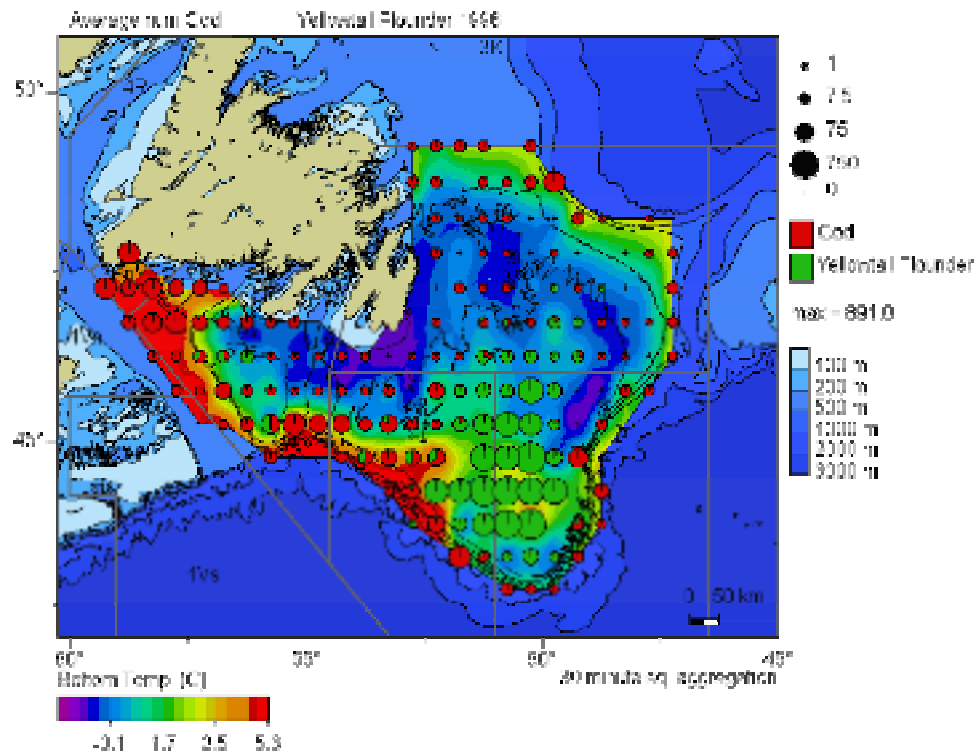


Fig. 5. Map of cod and yellowtail flounder bycatch, denoted by scaled symbol pie charts, in the American plaice fishery. Coloured surface represents bottom temperature created using inverse distance weighting of bottom temperature measurements.

The use of ACON as an interactive web mapping tool was also demonstrated. The web site:

<http://gmbis.marinebiodiversity.ca/>

utilizes ACON as a CGI (Common Gateway Interface) program that allows the user to interactively generate maps of data from the Gulf of Maine groundfish surveys and the Atlantic Reference Centre (ARC) specimen collection. The aim of the Gulf of Maine Biogeographic Information System project (GMBIS) is to develop, test, and demonstrate data dissemination and visualization tools involving overlaying numerous environmental and ecosystem data layers in the Gulf of Maine region. The project explores how a GIS and Internet technologies can be used to access databases and display their products for addressing specific questions relating to the biogeography and status of marine populations.

As part of this web site, the browsing of the “Electronic Atlas of Specimens from the Atlantic Reference Centre of the Huntsman Marine Science Centre” was presented (Fig. 6). In this page, 124000 specimens from the Reference Centre can be browsed at the Order, Family, or Genus level, providing a synopsis of the significant species retrieved for a specific mapped region.

A significant feature of the GMBIS web site is a “Map your own Data” web page, in which users can map their own data by uploading ASCII data files containing data fields such as longitude, latitude, and the mapping variable values to the web site. Scientists wanting to visualize their own data without investing in their own GIS technology, are able to generate their own maps of the NAFO region by uploading their own data to this site. A variety of output formats are provided.

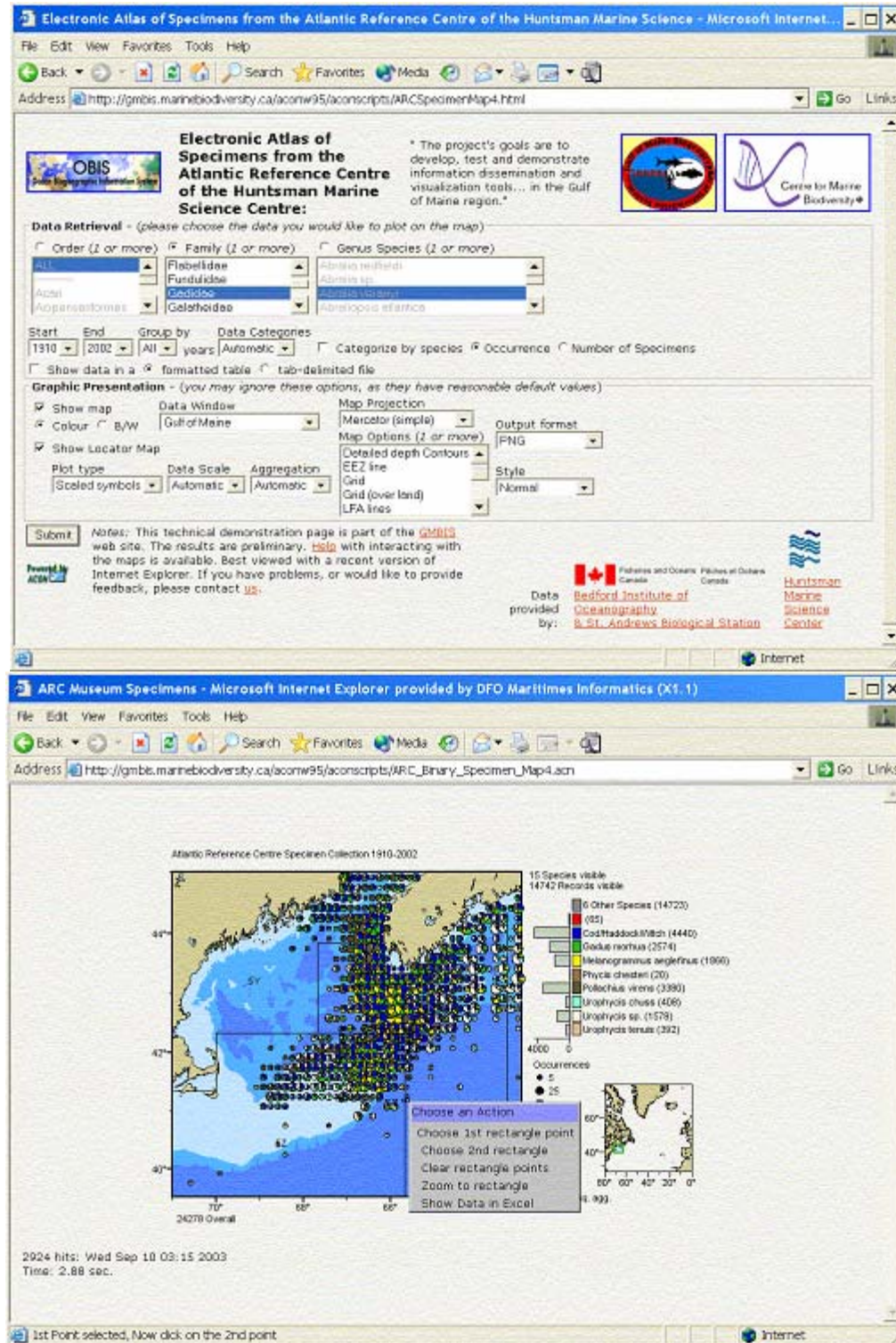


Fig. 6. (Upper panel) GMBIS selection page for interactive mapping of specimens from the Atlantic Reference Centre of the Huntsman Marine Science Centre. (Lower panel) GMBIS web page for the interactive electronic atlas showing output map.

3.1.2 SPANS

D. Kulka illustrated some of the capabilities of SPANS, GIS/geostatistical software with import and export routines that facilitate compatibility with other GIS and graphics software. Although not freeware, it contains point to surface functions including potential mapping routines plus overlay modeling capabilities that are particularly well-suited for fisheries data analysis. Various point to surface results were reviewed including point aggregation, Voronoi, Contour (TIN) and Potential Map (Fig. 7). Refer to Anon (1999) for further details on the application of each of the techniques.

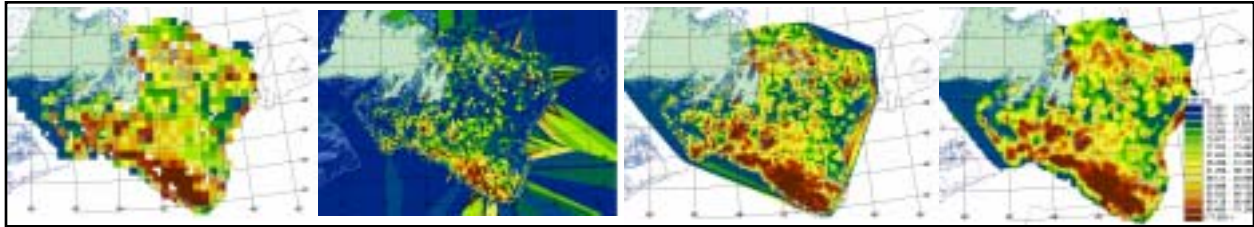


Fig. 7. From left to right, an example of the same survey data displayed as a point aggregation, Voronoi polygons, Contour (TIN) and a potential map derived from point data indicating survey catches (kg per tow) of American plaice on the Grand Banks.

The advantages and short comings of each technique was discussed. The potential map has a number of advantages over the interpolation methods. It does not extrapolate beyond the influence of the point data upon which it is based as is the case for the Voronoi and Contour (TIN) surfaces. It also has an array of functions that allow the user to control the output.

Potential mapping (far right panel, Fig. 7) is most appropriate for interval or ratio point data that represents a non-continuous variable, typified by a high degree of variance and contagious distribution, such as fish and fishery distribution data. Potential mapping converts point data, such as fishing sets, to a surface representative of a selected mapping variable or attribute (Z-value). Functions include: density of points (number of sets per km²), weighted average of an attribute such as catch rate or environmental variables, standard error of the mean, plus other measures of variance. Potential mapping uses an averaging technique and does not result in the creation of any new values outside the range of the input data and does not result in extrapolation beyond the influence of the original data. The model for weighted values of Z is defined by the equation in Fig. 8.

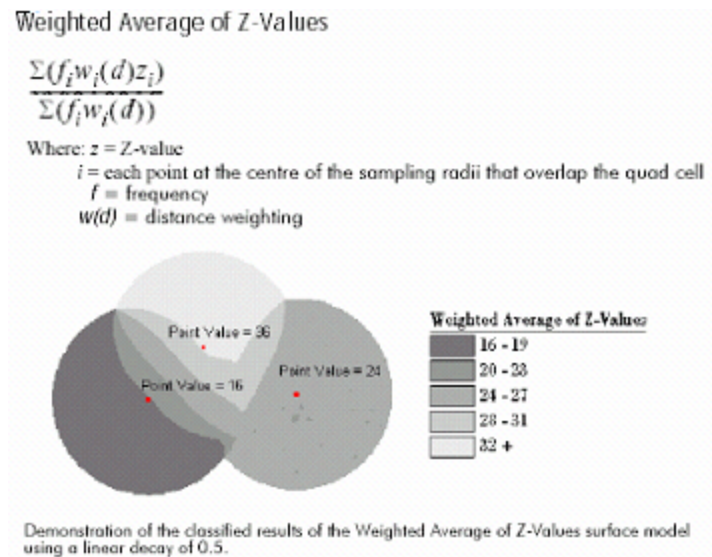


Fig. 8. Potential mapping model for the weighted value of Z (attribute associated with a geo-reference) (from Anon 1999).

The potential mapping derivative function generates a surface by applying a sampling radius to each point in the data layer (Fig. 9, left panel).

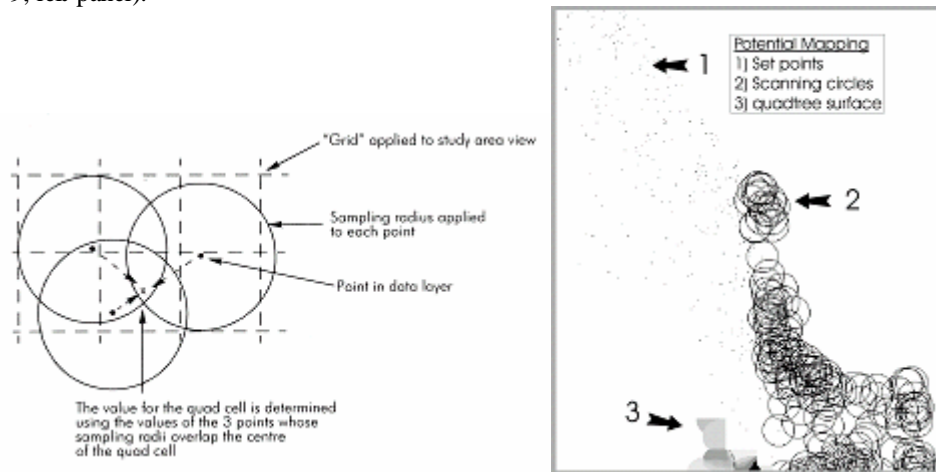


Fig. 9. Creation of a potential map from points using a circle of influence (left panel after Anon, 1999). The right panel shows: 1) the data points; 2) the resulting crescents, each with a unique value, formed by the overlap of circles; and 3) the user-classified surface derived from the crescent values.

This process is performed on all of the points and effectively creates a very large number of crescents or circle fragments. The value of these crescents are assigned to an underlying grid. No output values are calculated for areas lying outside of any sampling radius.

When creating a surface, the user can control the effect of distance on the resulting output values by specifying the size of the sampling radius and a rate of decay for decreasing the influence of points from the center. A classification scheme is applied to the output. Each of these areas has a unique value that is assigned to the underlying grid and these entities can then be classified by the user into a continuous surface describing an attribute such as a catch rate or environmental variable.

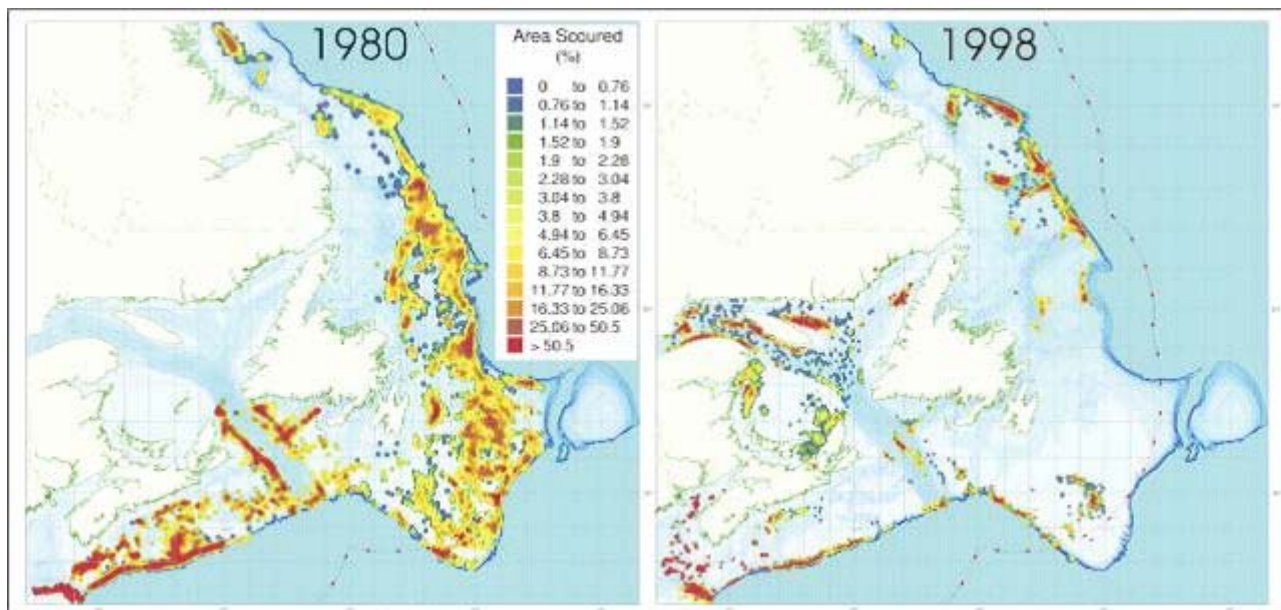


Fig. 10. Maps of trawling intensity in the Canadian Atlantic, during 1980 and 1998, created using the density function of potential mapping in SPANS (after Kulka and Pitcher 2001).

Practical applications of this method were presented using the results from previous studies. For example, the density function was used to create a series of annual maps depicting intensity of trawling represented as percent of area trawled (Fig. 10). This yielded a precise spatial definition of where the affects of trawling would be greatest.

In order to illustrate how a point to surface overlay can be used to solve a practical fisheries problem, the trawling intensity surface was overlain on points indicating juvenile cod distribution. This spatial approach effectively indicated that the overlap between 0-3 year-olds and moderate to heavily-trawled areas was minimal (Fig. 11).

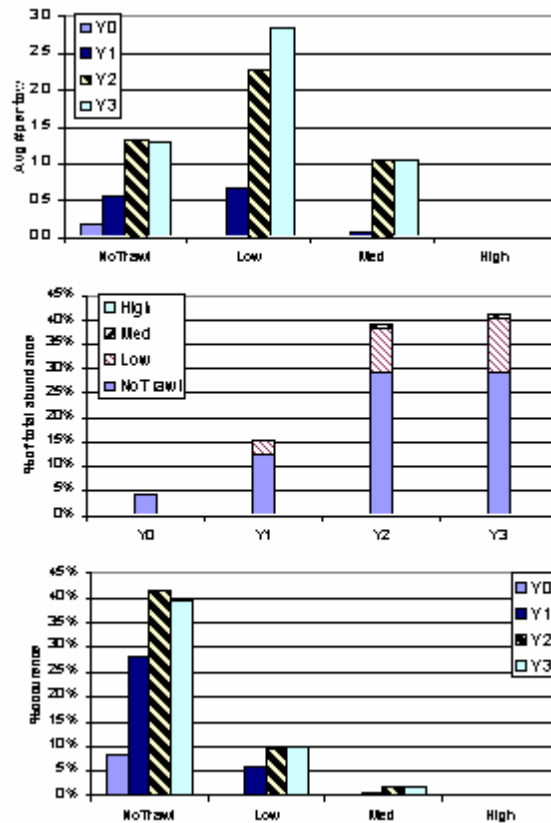


Fig. 11. Average number per tow (top), percent of total abundance (middle) and percent occurrence (bottom) of 0 to 3 year-old cod in non-trawled to highly-trawled areas.

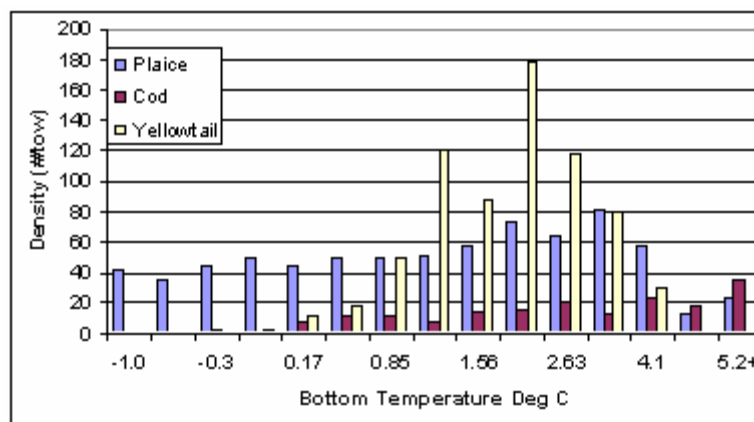


Fig. 12. Relationships between American plaice, cod and yellowtail flounder densities (number per tow) and bottom temperatures, derived by a point (species density) to surface (temperature) overlay.

To further demonstrate analytical capabilities of the software, survey catches of American plaice, yellowtail flounder and cod were overlain on a temperature surface, created by potential mapping, to define environmental relationships (Fig. 12). An advantage is that data from multiple sources can be analysed together in a spatial environment.

A second example demonstrated how potential mapping can be employed to estimate biomass and abundance using either survey or commercial fisheries data (Kulka and Pitcher1998). Using the example of thorny skate, potential mapping was used to post-stratify the data based on the distribution of the species. Areal expansion was then applied within the density strata created by potential mapping (Table 1), and because the strata correspond to the distribution of the species, the result is fewer strata with more stations per stratum. Data from any type of survey or even commercial fisheries data can be analysed in this manner.

Table 1. Calculation of thorny skate biomass based on density classes (strata) created using potential mapping (refer to Kulka and Pitcher 2001 for details).

Density Class	Set Count	3LNOP Area	Mean Kg per tow	Biomass kg	Biomass t	Stdev
1	21	25,051	0.0	0	0	0.0
2	26	24,750	0.1	105,924	106	0.2
3	29	25,256	0.4	421,548	422	0.4
4	28	27,236	1.0	1,142,318	1,142	1.0
5	26	23,974	0.9	943,943	944	1.1
6	27	23,519	1.4	1,419,239	1,419	1.7
7	35	23,979	2.5	2,606,647	2,607	2.8
8	21	22,660	3.5	3,384,565	3,385	3.5
9	33	19,122	5.8	4,713,823	4,714	5.1
10	28	19,138	7.9	6,437,787	6,438	7.5
11	35	18,356	9.6	7,541,657	7,542	8.3
12	31	18,781	15.5	12,458,561	12,459	17.8
13	29	21,666	23.5	21,762,527	21,763	19.9
14	30	22,070	30.0	28,364,501	28,365	22.2
15	26	16,402	53.4	37,456,752	37,457	43.5
Sum	425	331,960	10.36	128,759,791	128,760	9.00

3.1.3 Ocean Data View

R. Schlitzer demonstrated the capabilities of Ocean Data View (ODV), freeware available on the Internet at <http://www.awi-bremerhaven.de/GEO/ODV/>, which was designed for the interactive exploration and graphical display of multi-parameter profile or sequence data. The software is run on PCs under Windows and Linux, on Macintosh under OS X and on UNIX workstations under SUN Solaris, SGI IRIX, and IBM AIX. The software includes coastline and topographic data as well as various databases with official names for a large number of topographic features (gazetteers). Although originally developed for oceanographic observations, the underlying concept is more general and data or model output from other disciplines such as fisheries research can be achieved with ODV. The data format is designed for compact storage as well as direct data access and allows the construction of very large datasets. ODV supports the display of point data as either colored circles or numeric data values. In addition, variable-resolution gridding algorithms allow color shading and contouring of gridded fields along sections and on general three-dimensional surfaces. A large number of derived variables can also be selected, calculated and displayed on-line.

3.1.3.1 Operational Modes

ODV can operate in five different modes (MAP, STATION, SCATTER, SECTION and SURFACE), thereby providing different analysis methods and display types commonly used in the scientific community. MAP mode can be used to produce high quality station or cruise maps of specific regions or the entire globe. ODV allows a choice between five map projections and provides bathymetry and land topography information as well as the boundaries of rivers, lakes, sea-ice extent and national borders. This information is available at different levels of resolution and

can be used to compose various types of context maps. Specific stations can be highlighted and annotated. ODV maps can be produced as GIF, EMF, or PostScript files.

STATION mode (and all of the following modes) provides a station map and one or more data plot windows. This mode is appropriate for producing X/Y property/property plots for one or more selected stations or profiles and for studying differences between stations. The stations to be plotted can be selected using different methods, the simplest method is to click on a station with the mouse. In SCATTER mode, data from all stations shown on the map are displayed in the data plots. This provides an overview of all data from a given region, a specific cruise, or a selected station subset and is particularly useful for data quality checking. SCATTER mode (and all of the following modes) supports Z variables in addition to X and Y variables. The value of a Z variable at a given X/Y point is displayed as either the actual numerical value or by value-dependent color coding. Plots with Z-variables (similar to SECTION and SURFACE plots described below) can be displayed in two ways: (1) as colored circles at the X/Y locations or (2) as continuous, gridded fields estimated on the basis of the observed data. Gridded fields (see below for a description of the gridding algorithms) can be color-shaded and/or contoured.

SECTION mode also supports Z variables on data plots and allows all plot types of the SCATTER mode, but the set of stations used for the plots is restricted to a section band usually following given cruise tracks. Section bands can be defined arbitrarily and their width can be adjusted in order to properly select a set of stations. SECTION mode is appropriate for presenting property distributions and property/property plots for all stations along entire cruises and to calculate and investigate geostrophic velocities perpendicular to the cross-section of a cruise track.

SURFACE mode allows the specification of surfaces in three-dimensional space which are defined as points of constant value for a given variable. For example, depth, density, or temperature surfaces can be displayed and overlain on property distributions of other variables. Fig. 13 shows an example of SURFACE plots produced from fish catch data in the Grand Banks region.

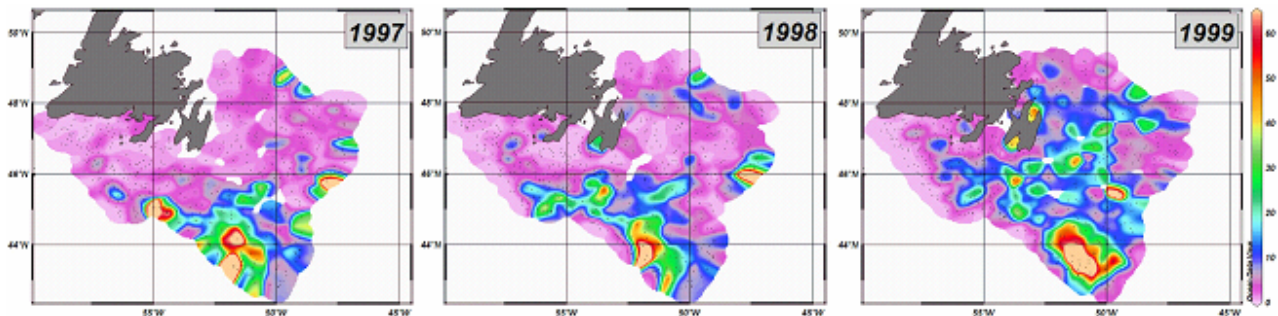


Fig. 13. Distribution of American plaice based on 1997-1999 survey catches (kg per tow) in the Grand Banks region as examples of ODV map displays. The black dots represent the positions of the original data.

3.1.3.2 Gridding Algorithms

In addition to displaying data points as numerical values or colored circles, ODV can produce color-shaded and/or contoured, gridded property distributions (Fig. 13). For the gridding process, ODV has two built-in gridding algorithms: *Quick Gridding* and *VG Gridding*. *Quick Gridding* is a fast method suitable for cases with good data coverage and yields results in a matter of seconds, even for large datasets that include as many as several hundred thousand points. The underlying method is a weighted-average scheme with separate user-supplied averaging length scales in X and Y directions. To achieve fast performance, the data are tile-sorted prior to the averaging process then, for the estimation at a given X/Y position, only the data values within a small neighborhood of the point are actually used in the averaging process.

For poor or heterogeneous data coverage, *VG Gridding* is to be preferred over the *Quick Gridding* method. In contrast to *Quick Gridding*, which uses an equidistant, rectangular grid for the estimation, *VG Gridding* analyzes the distribution of the data points and constructs a variable-resolution, rectangular grid, where grid spacing along the X

and Y axes vary according to data density. High resolution (small grid spacing) is provided in regions with good data coverage, whereas in areas of sparse sampling the grid is coarse and resolution is limited. For typical hydrographic sections, this procedure leads to higher spatial resolution in the upper water column and in boundary current regions (data coverage is usually very good in these areas) as compared to the deep, open ocean regions.

After construction of the grid, the property under consideration (e.g., temperature, salinity) is estimated at every grid point by applying a weighted-average scheme using data values from the grid point neighborhood. Weights decrease with increasing distance from the grid point, and user-specified length scales in X and Y directions are applied. Averaging length scales are proportional to the grid spacing. For example, in areas of higher grid-resolution such as the upper water column and boundary currents, smaller averaging length-scales are used automatically. This overall approach allows the resolution of small-scale features in areas of dense data coverage and provides smooth and stable fields in areas with sparse data coverage. Once a property field has been estimated, the results are passed to shading and contouring routines output as a screen display or as a printable file.

3.1.4 Generalized Additive Models

Modelling of the spatial distribution of fish populations using non-parametric Generalized Additive Models (GAM's) was presented by M. Simpson. It was noted that while trend surface analysis could be used to model the spatial distribution of catches, higher-order polynomial regressions provide poor fits along the edge of the spatial distribution and only capture global patterns of the distribution. In contrast, GAM's extend the range of Generalized Linear Models by allowing non-parametric smoothers. Employing GAM's, spatial trends in catch in relation to environmental can be investigated using the R software package.

Two smoothing functions are available in R: **s** (cubic B-spline) and **lo** (loess), which can be used on their own, or mixed with parametric functions. GAM's permit specification of the error distribution, such as binomial, gamma or Poisson. The latter type of error distribution is appropriate for data collected during research surveys (Swartzman *et al.*, 1992; O'Brien and Rago, 1996). The model is fitted by iteratively by smoothing partial residuals. The syntax involved in running the GAM in S-plus, a detailed S-plus code outlining a similar application of the GAM to fisheries data, available in O'Brien and Rago (1996, NAFO Sci. Coun. Studies 28:79-95), was followed in the presentation.

A step-wise GAM that incorporated yellowtail flounder survey catches (kg per tow) was used to determine the best fit based on the Akaike Information Criterion (AIC) test statistic. The lowest AIC statistic gives the best combination of parameters for inclusion in the final model. To define the test criteria, the GAM equation was specified as:

$$\text{gameq} < \text{formula}(\text{yelnum} \sim \text{s}(\text{DEPTH}) + \text{s}(\text{BTEMP}) + \text{LAT} + \text{LONG})$$

where yellowtail flounder catch, (yelnum), was modelled as a function of depth, bottom temperature, latitude and longitude. The step-wise criteria, in which the form of the smoother (i.e., lo or s) can be specified in a scope list for each environmental variable, was defined as follows:

$$\begin{aligned} \text{"DEPTH"} &= \sim 1 + \text{DEPTH} + \text{lo}(\text{DEPTH}) + \text{s}(\text{DEPTH}), \\ \text{"BTEMP"} &= \sim 1 + \text{BTEMP} + \text{lo}(\text{BTEMP}) + \text{s}(\text{BTEMP}), \end{aligned}$$

The step-wise routine generated an AIC statistic for each iteration to evaluate the model.

Start: $\text{yelnum} \sim \text{s}(\text{DEPTH}) + \text{s}(\text{BTEMP}) + \text{LAT} + \text{LONG}$; AIC= 13960.87

Trial: $\text{yelnum} \sim \text{lo}(\text{DEPTH}) + \text{s}(\text{BTEMP}) + \text{LAT} + \text{LONG}$; AIC= 15791.23

Trial: $\text{yelnum} \sim \text{s}(\text{DEPTH}) + \text{lo}(\text{BTEMP}) + \text{LAT} + \text{LONG}$; AIC= 14469.87

Trial: $\text{yelnum} \sim \text{s}(\text{DEPTH}) + \text{s}(\text{BTEMP}) + 1 + \text{LONG}$; AIC= 14729.04

Trial: $\text{yelnum} \sim \text{s}(\text{DEPTH}) + \text{s}(\text{BTEMP}) + \text{LAT} + 1$; AIC= 16656.19

An evaluation of the model results demonstrated that the cubic B-spline smoother provided the best fit. Following model verification, the GAM produced a plot of the interpolated values in the survey area by using the commands:

```

yellowtailGAM<-gam(gameq,family=poisson,data=yellowtail)
i<-interp(LONG,LAT,yellowtailGAM$fitted)
image(i,xlab="Longitude",ylab="Latitude",xlim=c(-55,-47))
contour(i,add=T,nlevels=15)
zcat<- (yellowtailGAM$fitted)

```

The commands shown above will produce a plot of the fitted output from the GAM which can then be annotated with additional commands to specify the title, legend, coast, depth contours, and NAFO Division boundaries (Fig. 14).

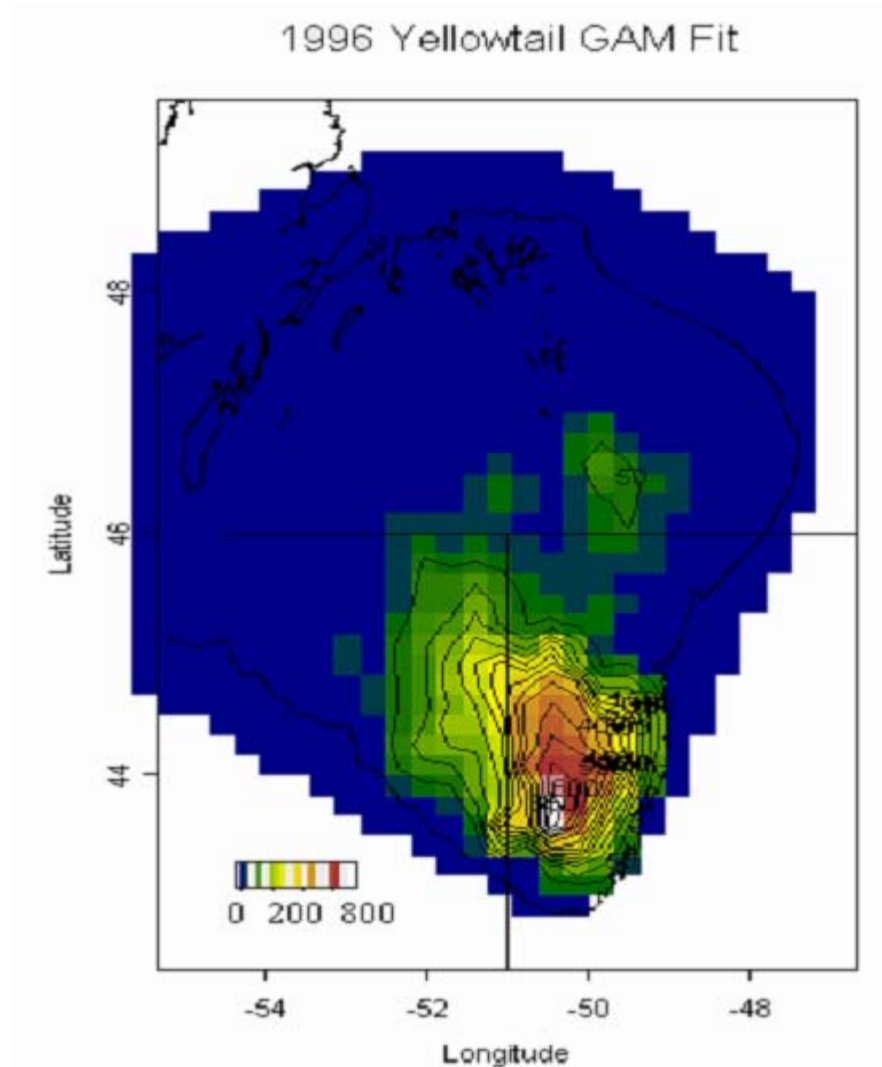


Fig. 14. Map of yellowtail flounder distribution on the Grand Banks based on spring survey data fit using a GAM.

To evaluate the model, the model summary can be produced by issuing the commands `summary(yellowtailGAM)` to display the Deviance Residuals:

```

Null Deviance: 86140.62 on 503 degrees of freedom
Residual Deviance: 13938.91 on 493.022 degrees of freedom
Number of Local Scoring Iterations: 6

```

from which a Pseudo- R^2 (O'Brien and Rago, 1996) can be calculated using the formula $(1 - \text{residual deviance} / \text{Null deviance})$. In this example, the Pseudo- R^2 is 0.8381 and provides an indication of the model fit.

GAMs, in contrast to some analytical procedures, do not make *a priori* assumptions about underlying data relationships. As a result, the data drive the fit of the model. The GAM also permits the user to visualize the additive contribution of each variable to the respective response using smoothed functions (Fig. 15).

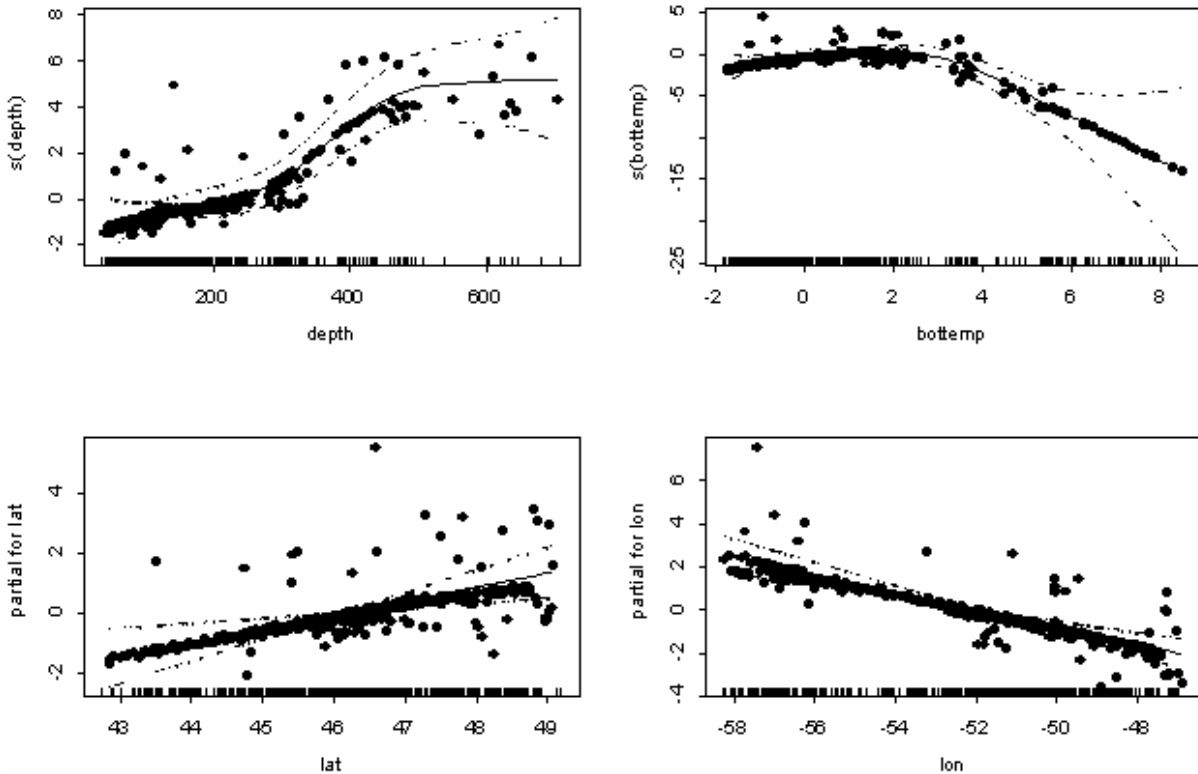


Fig. 15. Scatterplot smooths and partial fits, with 95% confidence limits (dashed lines), of depth (top left), bottom temperature (top right), latitude (bottom left) and longitude (bottom right) for a Generalized Additive Model (GAM) of yellowtail flounder spring survey catches (kg per tow) on the Grand Banks.

GAMs constitute a powerful exploratory tool for detecting simple and complex relationships in survey distributions. Unlike tessellation methods, in which exterior polygons/triangles are undefined and no variance is associated with the interpolated points, or trend surface analysis, which only shows global scale trends and has difficulty fitting the surface near edge of the distribution. Generalized additive models are better suited to handle the nonlinear relationships between fishery catch distributions and associated environmental variables (e.g., depth, temperature, sediment type, and salinity).

3.2 Demonstration - Solving a Spatial Question

Practical examples of how spatial analysis can be used to solve a fisheries question were demonstrated in ODV, R, ACON and SPANS using the Grand Banks data. The objective of the analysis was to define the overlap of two species in terms of their geographic distributions.

3.2.1 Ocean Data View

To determine the geographical regions, where the two species American plaice and yellowtail flounder overlap, ODV was used to generate the distribution maps for the two species based on the multi-year Grand Banks dataset (Fig. 16a and b).

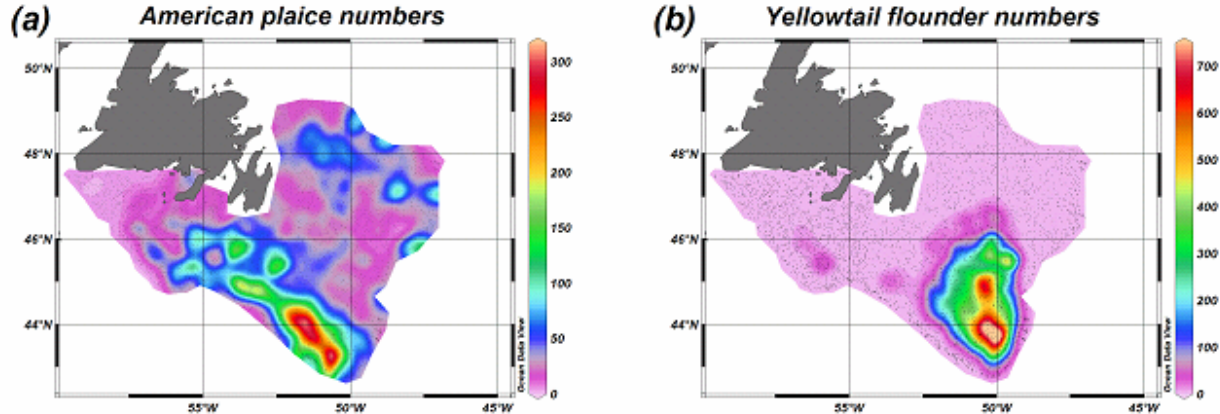


Fig. 16. Distribution of (a) American plaice and (b) yellowtail flounder in the Grand Banks region for the years between 1996 and 2003

It is evident that American plaice has its highest abundance in a northwest-southeastward oriented band in the southern Grand Banks area, whereas yellowtail flounder is most abundant in a meridionally stretched region somewhat to the east of the American plaice patch. There appears to be overlap of the two species, however, the precise location and extent of the overlap area is difficult to determine from Fig. 16 alone.

For a quantitative treatment of the problem the American plaice data are plotted versus yellowtail flounder (Fig. 17a), and the data points in the parameter range that indicates overlap (simultaneous high yield for both species; marked by a red polygon in Fig. 17a) are highlighted in green. When plotted in a map (Fig. 17b), one finds that the green points form a well defined area of overlap in the southern Grand Banks region. Note that the definition and highlighting of overlap points was implemented by using ODV's Patch derived variables. Refinements that also show the degree of overlap would be possible by specifying more than one patch.

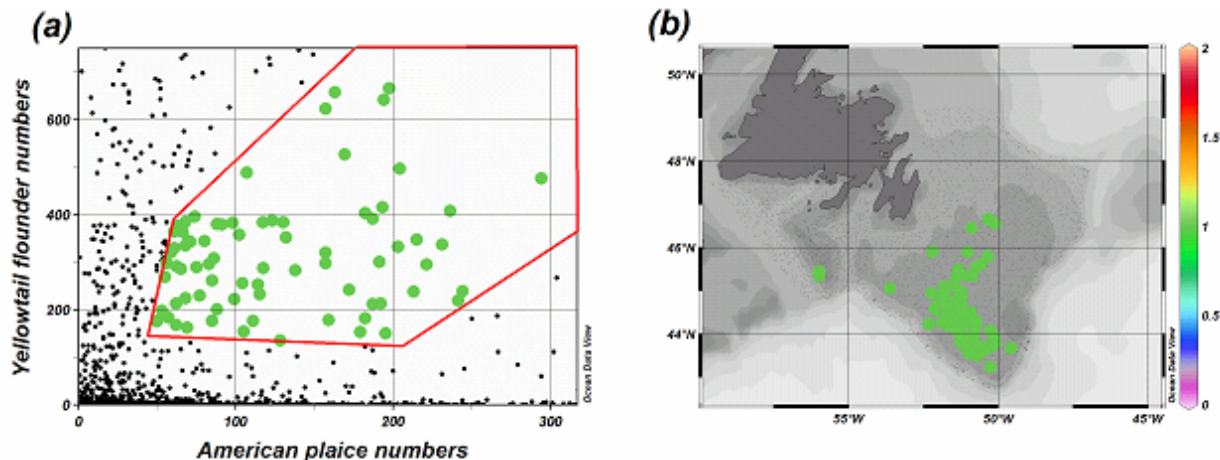


Fig. 17. (a) Plot of yellowtail flounder versus American plaice and specification of a parameter range that indicates overlap (red polygon and green data points); and (b) map of the Grand Banks region showing the geographical distribution of overlap points.

3.2.2 R-Geostatistics

R-geostatistical routines were applied to quantify the spatial overlap between yellow tail flounder and the American plaice. Following Bez and Rivoirard (2000), each population is represented by an ellipse centered at the center of mass of the population. As noted later in this report, the inertia of the mass of fish around their central location, i.e. the fish dispersal, can be decomposed into the directions in which fish are most and least dispersed, which results in defining a fish distribution as an ellipse. The surfaces of the ellipses are equal to the equivalent surface (Bez and Rivoirard, 2001) of each population (Fig. 18). Finally, the spatial overlap between the two can be quantified by a Global Index of Collocation which ranges from 0 when the overlap is null to 1 for complete overlap. In the present case, the overlap appears to be low as the GIC remains below 0.8. In the meantime, a trend is observed indicating that the spatial overlap between the two species increased during the study time window (Fig. 18).

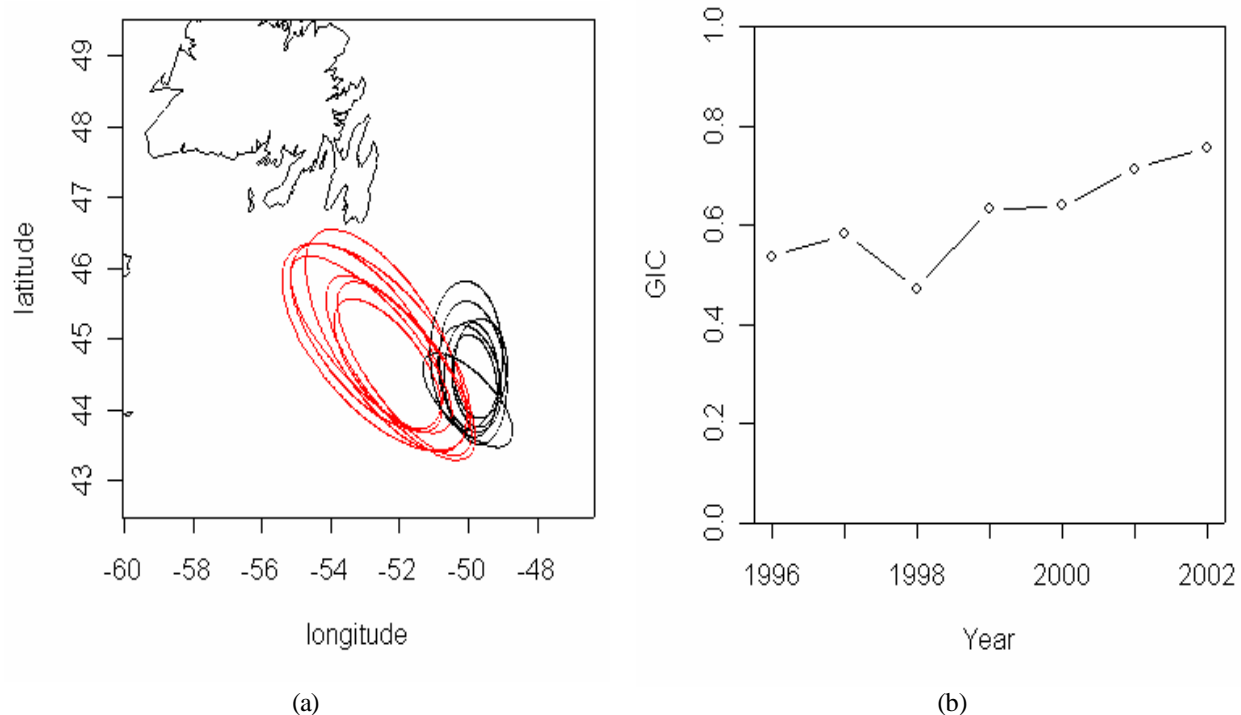


Fig. 18. Evaluation of the spatial overlap between yellowtail flounder (black ellipses) and the American plaice (red ellipses) for the different surveys. (a) Summary spatial distributions using centres of mass and inertia. (b) Evolution of the Global Index of Collocation (GIC) with time.

3.2.3 ACON

A geo-referenced pie chart graphic of the co-occurrence of both yellowtail flounder and American plaice was generated to define both the distribution and co-occurrence of these species and their relative abundance (Fig. 19). In this example, aggregation by 10 minute square provides a description of fine scale structure along the shelf edge in areas of high sampling intensity. Although in this approach there is no error introduced as may be the case for alternative contouring methodology assumptions, no global estimate is available for either the frequency, magnitude or variance of co-occurrence. Bottom temperature observations are contoured in this graphic, as it is traditionally considered a continuous variable for which one assumes interpolation between adjacent observations produces a reasonable representation. However, this script included a parameter that limits the spatial extent of contour interpolation, so that the Delaunay triangulated bottom temperature data in the upper right corner of the map were determined to extent the interpolation beyond an acceptable (arbitrary) sampling distance. As a result a “hole” appears in the contour surface.

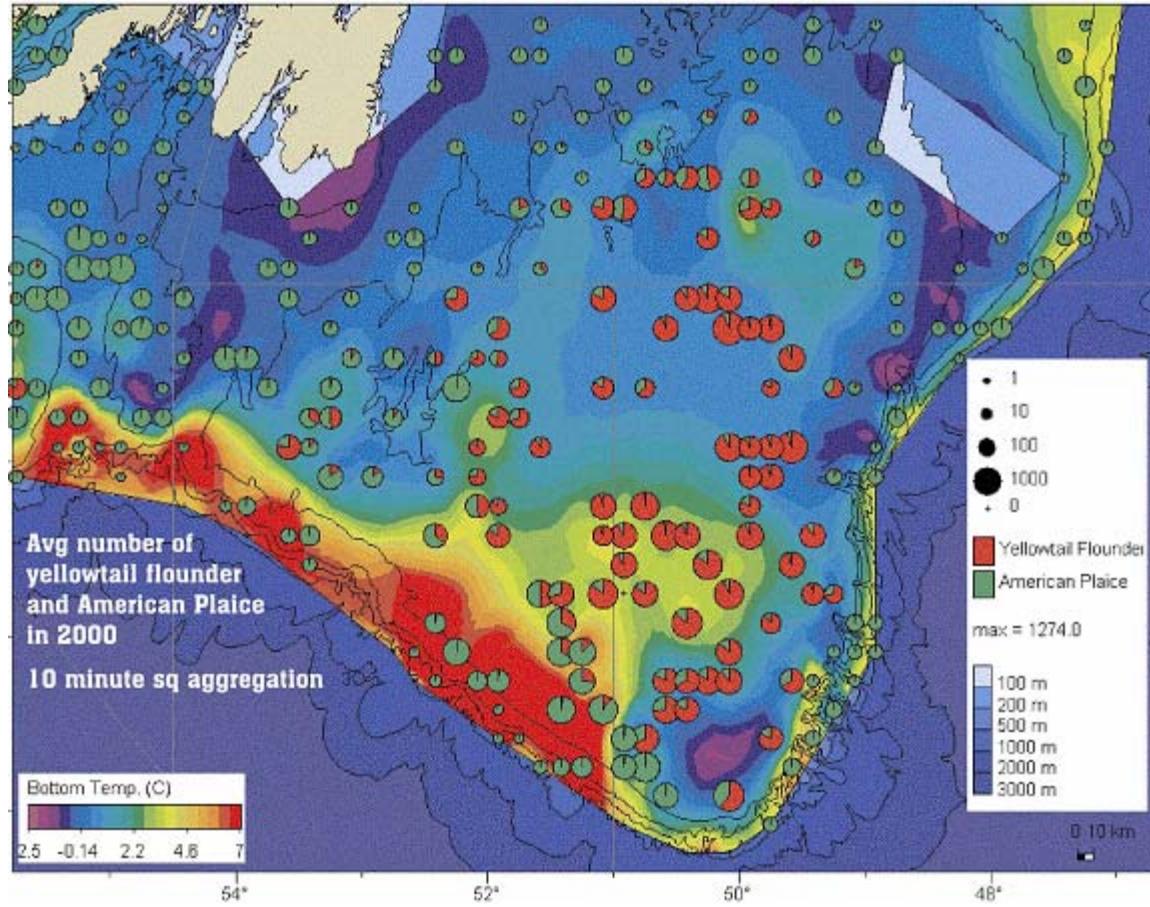


Fig. 19. ACON generated map of Yellowtail Flounder and American Plaice, denoted by scaled symbol pie charts, in the American plaice fishery in 2000. Coloured surface represents bottom temperature created using inverse distance weighted gradient interpolation of bottom temperature measurements.

3.2.4 SPANS

Matrix modeling was applied to spatially define overlap of the distributions, potential maps of American plaice and yellowtail flounder were created from the survey data (kg per tow). The two maps were then overlaid using a matrix model to produce a matrix of zeroes where columns defined yellowtail density (kg per tow) and the rows defined American plaice density. The matrix was then revised using a numerical classification scheme that reflected the degree of spatial overlap between the two species (table in Fig. 20). This matrix classification was then used to produce a map showing the degree of overlap between the two species (map in Fig. 20).

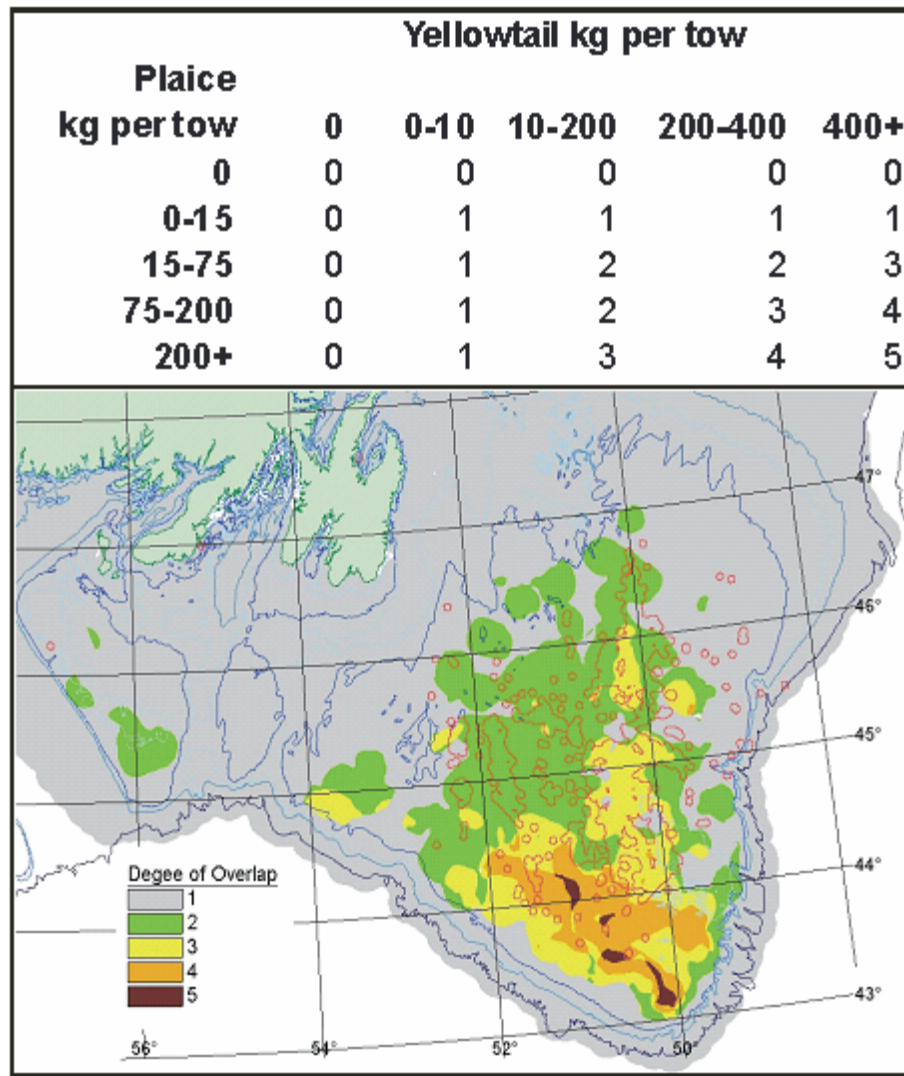


Fig. 20. Matrix overlay analysis of American plaice and yellowtail flounder. The table above the map shows the matrix used to reclassify the overlay of two potential maps. The map resulting from the matrix overlay shows the degree of overlap between the two species (1-no overlap, 5 greatest overlap). These areas of overlap can then be compared to fishing grounds (red lines overlaying the area).

4.0 UTILIZATION OF GEOSTATISTICS FOR MAPPING

N. Bez presented a lecture on the utilization of geostatistics for mapping. Lectures pertaining to geostatistical concepts and methods were followed by hands-on exercises that incorporated survey and fisheries data from the Grand Banks.

4.1 History, theory and concepts

A brief introduction was given on the origin of geostatistical techniques in the mining industry to solve systematic over-estimation of gold reserves. The use of geostatistical tools in fisheries has gradually increased since 1977 when Laurec (1977) first used geostatistics to estimate fishing power. Thereafter, Conan (1985), Laloé (1985), and Gohin (1985) utilized geostatistics to analyze fisheries data. Since 1985, fisheries applications of geostatistical methods have focused primarily on the analysis of acoustic survey data, in part, because its use was recommended as an outcome of a 1991 ICES workshop on the topic (ICES, 1993).

4.2 Consideration of the spatial aspects of fish distributions

The classical method of estimating the quality of estimations is to compute the estimation variance by using the formula:

$$\sigma^2_E = \frac{s^2}{N}$$

where N represents the number of samples, s^2 represents the experimental variance of the data and σ^2_E represents the estimation variance. However, the equation is based on the assumption that the number of samples can be considered as N outcomes of N variables that are independent and identically-distributed in space (spatial homogeneity of the distribution). Independence can be attained in practice if the samples are located at random or if the underlying fish distribution has no spatial structure, but this is seldom the case.

When the objective of a survey sampling scheme is either to target a particular population or to sample the distribution of multiple species, the area of occupancy of a given species is often bounded by stations with no catch of the given species. These zeros represent valuable information because they indicate whether the area of occupancy of the target species has been adequately sampled. However, for the same reason, zero catches also impact the classic tools used for the analysis of survey data, such as histograms, means, variances, and regressions. The alternatives are either to utilize tools that are robust with respect to zero catches (e.g. transitive geostatistics) or to delineate the area occupied by a species or variable then to statistically describe the characteristics that are intrinsic to each population. This latter approach, intrinsic geostatistics, consists of the use of variograms and the production of kriging maps.

4.3 Identification of spatial structure using variograms and kriging

The variogram is a tool for describing the spatial structure of a variable. The variogram was presented as the decomposition of the overall variance into distance classes. The variogram allows for the determination of which spatial scales (small or large-scale) are most responsible for the overall variance. An example of a one-dimensional variogram was presented and participants computed the variogram, by hand, in order to better understand application of the algorithm. Practical considerations for the estimation of two-dimensional (bi-directional) variograms were also presented. Experimental variograms are sensitive to: distance lag, field size and shape, outliers and their location in the field, and the homogeneity of the population.

The effects of adjusting the distance and direction lags on the variogram were considered. The possible use of tolerance values, when samples are not regularly spaced, was discussed and the avoidance of overlapping distance and direction classes was recommended to avoid double counting.

An example based on the piezometric levels of an aquifer at various geological depths was presented in order to illustrate the interpretation of experimental variograms. The example demonstrated that the evolution of the experimental variogram was clearly associated with the evolution from a pure nugget effect to a strong, small-scale spatial structure that increased with depth.

The variogram allows for the computation of the estimation variance. As a result, variogram models that preclude negative variance estimates must be chosen. Possible functions for use in fitting an experimental variogram were presented. Fitting rules available in the software include manual, semi-automatic and completely automatic. The decision for selecting a specific type of variogram model should be based on both the spatial distribution of the sample data and on biological knowledge of the small-scale spatial structure, as well as the level of measurement error associated with the survey gear.

The variogram model is used in kriging. Kriging allows for estimation of the density at an unknown point and uses an algorithm that defines the weights to apply to sample data when interpolating between known points in order to minimize the estimation variance. The mathematical basis of kriging was explained and the impact of the model parameters (type of model, range, or sill) on the kriging outputs was demonstrated using various examples.

4.4 Methods for summarizing spatial distributions

In some cases, particularly when the number of surveys and/or the number of species being studied is large, spatial distributions may be summarized using quick and more robust tools. For example, the center of mass and the inertia of sample values can be used to represent the mean location of a population and its dispersal around the mean position. Similar to a PCA analysis, the inertia can be decomposed into the directions in which fish are most and least dispersed, which results in defining a fish distribution as an ellipse.

4.5 Other geostatistical methods

Geostatistical methods other than kriging were only covered in brief. For example, particular attention was devoted to the difference between a local estimation (i.e. a kriging map) and a global estimation (i.e. the estimation of the fish densities over a given stratum, multiple strata, or a given area). In the case of a stratified, random survey, a kriging map can be used to estimate the biomass within a stratum by summing the kriging values estimated at the grid nodes that fall within the particular stratum multiplied by the area of a grid cell. However, a global estimation variance for each stratum must be computed using a different method (Rivoirard et al., 2001) and then the stratum variances can be summed across all strata to compute a weighted average.

Multivariate geostatistics and the use of cross variograms for co-kriging was mentioned briefly, as were the questions of species-specific survey design and the impact of the size of the area swept by the trawl.

4.6 Geostatistical analyses of Grand Banks data sets

Geostatistical analyses incorporating data sets from the Grand Banks were conducted using R software (freeware) and geostatistical routines developed by the Centre de Géostatistique. This allowed the participants to use and test the geostatistical tools presented.

In the case of multi-species surveys, the most likely situation is that the models have to be species and year-specific, but this is dependent on the objective of the analysis. Species present on the Grand Bank do not have the same areas of occupancy, so due to time limitations, analyses were focused on yellowtail flounder.

The spatial structure of yellowtail flounder was also compared to that of environmental variables, such as depth and bottom temperature, and biological variables such as Greenland halibut. The potential use of CPUE data to describe the spatial structure of targeted species was also examined.

4.6.1 Yellowtail flounder

Yellowtail flounder were concentrated in the southern portion of the Grand Banks during 1996-2002 (Fig. 21).

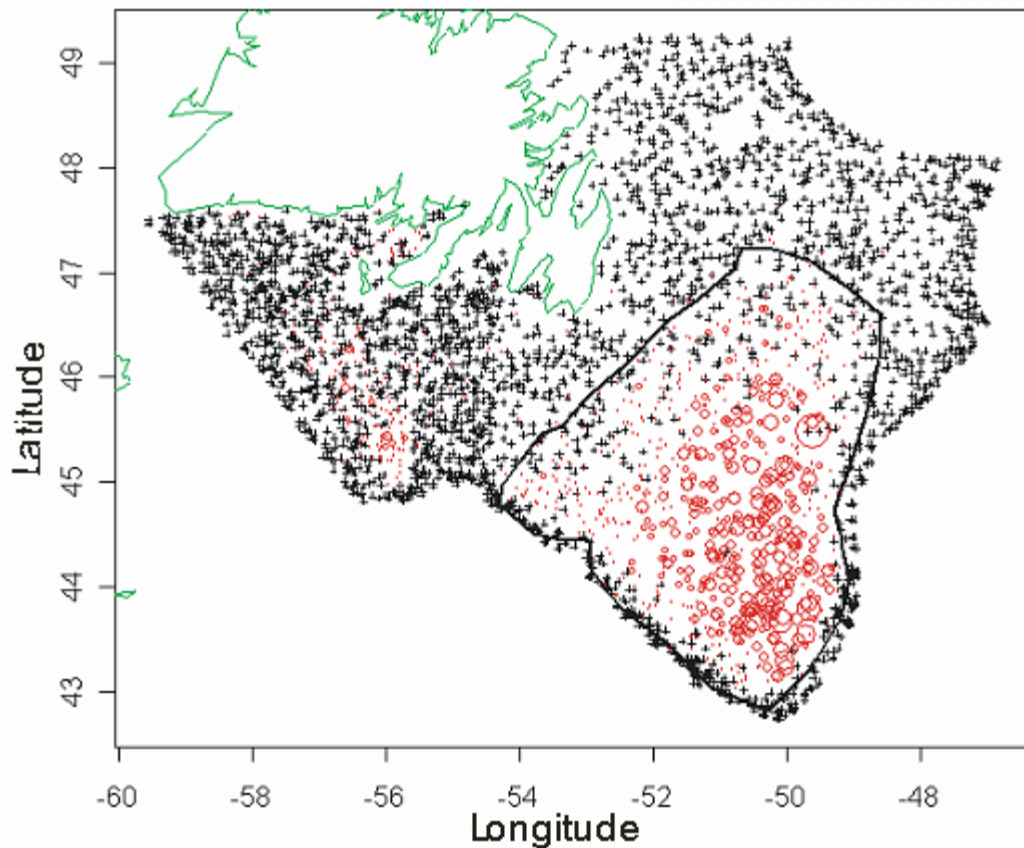


Fig. 21. Yellowtail flounder distribution (number per tow) during spring, based on the Canadian bottom trawl surveys conducted on the Grand Banks during 1996-2002, and the habitat polygon utilized in a kriging analysis. The diameter of the red circles increases with flounder density and the black crosses represent stations with no flounder catch.

Figures 22-24 illustrate the steps required to prepare a kriging map of yellowtail flounder abundance, during spring of 2000 ($N = 101$ stations), using the area of occupancy shown in Fig. 21. The selection of a particular area of occupancy affects the variance level and relative shape of the variogram (Fig. 22). As is often the case with fisheries data, the spatial structure of yellowtail flounder during spring of 2000 exhibited a strong random component (nugget effect). The nugget effect quantifies the amount of spatial structure that is unknown because it is attributable to inter-sample distance (the average distance between nearest neighbors) and measurement error. For the yellowtail flounder example presented, the nugget effect represents half of the overall variability. The rest of the variability is explained by a spatial structure with a range of 80 nautical miles, which implies that fish densities at stations greater than 80 nmi apart are no longer correlated.

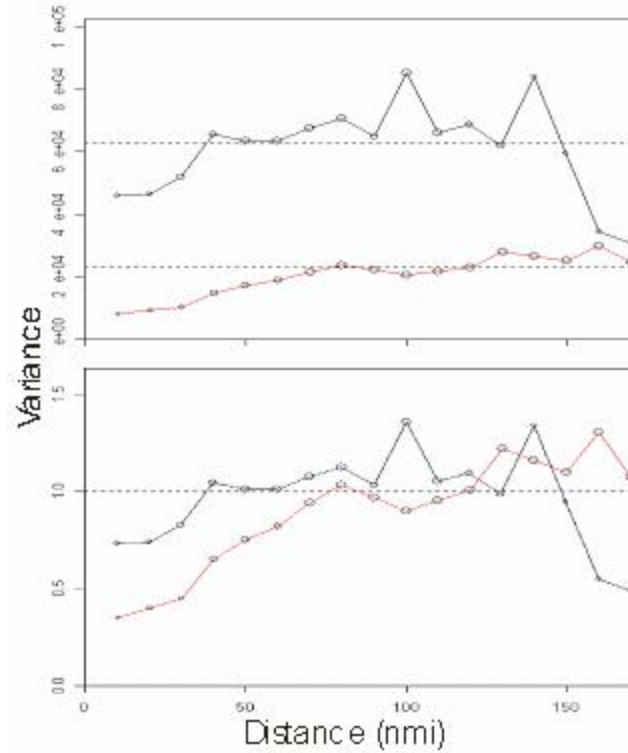


Fig. 22. Experimental variograms, presented in raw (top) and normalized variance scales (bottom), of yellowtail flounder abundance (number per tow) on the Grand Banks during 2000 based on Canadian spring bottom trawl surveys. Variograms are presented with (black line) and without (red line) the inclusion of stations with no yellowtail flounder catch and which lie beyond the boundary of the habitat polygon (refer to Fig. X1).

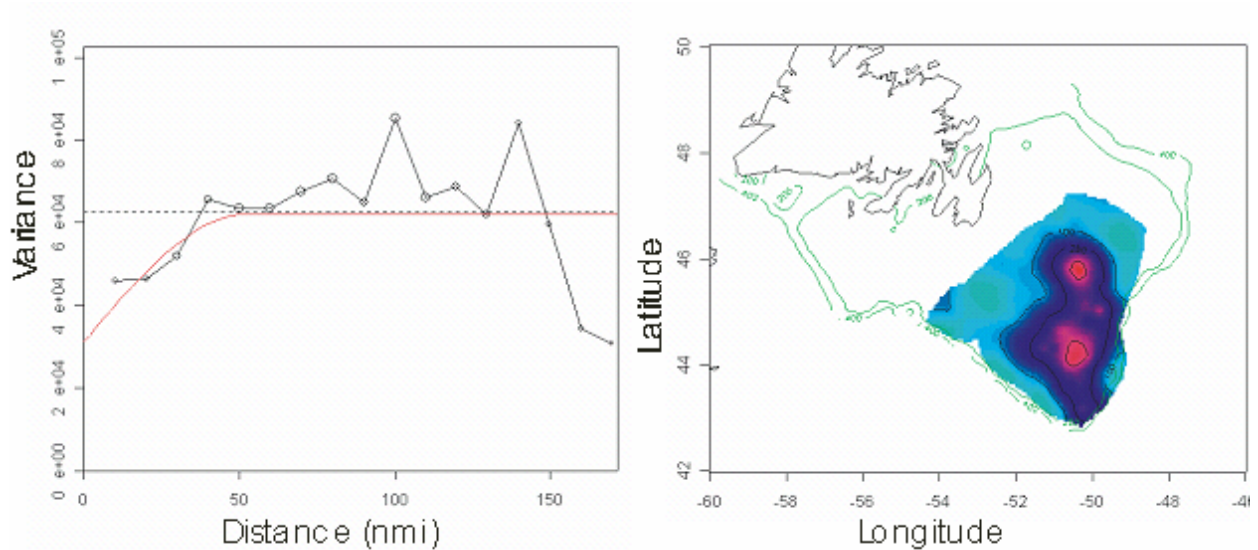


Fig. 23. An experimental variogram (left), computed for all directions, and variogram model (red line) for yellowtail flounder abundance (number per tow) on the Grand Banks based on data from the 2000 spring Canadian bottom trawl survey. The variogram and model are based on the habitat area delineated in Fig. X1. A kriging map (right) based on the variogram model is shown with iso-density contours (highest densities are shown in red) and the 200 m and 400 m isobaths are shown in green.

After a model was fitted to the experimental variogram (Fig. 23), a kriging map and a kriging variance map (Fig. 24) were prepared for the area of occupancy. The raw values for yellowtail flounder abundance were superimposed on both maps to check the accuracy of the interpolations.

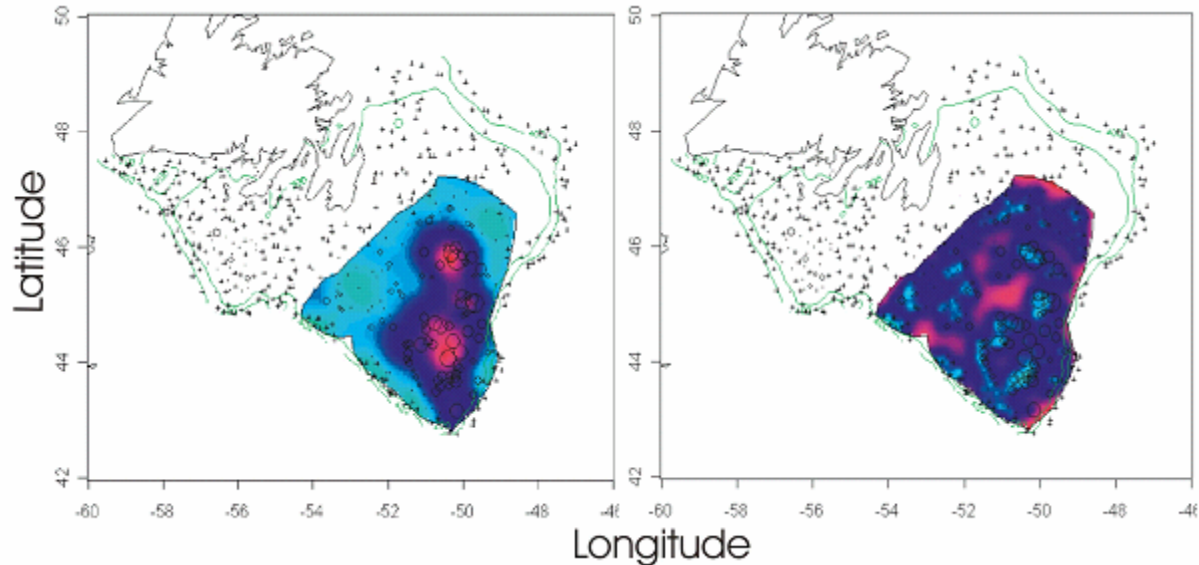


Fig. 24. Kriging map (left) showing the estimated distribution of yellowtail flounder on the Grand Banks, including the raw data points, and the associated kriging variance map (right). The highest densities and variance values are shown in red and the 200 m and 400 m isobaths are shown in green.

4.6.2 Depth

An experimental variogram for depth, a variable with, a priori, a strong spatial structure, can be compared with that of yellowtail flounder. The depth experimental variogram is more stable (Fig. 25) than the one obtained for yellowtail flounder. Nevertheless, depth exhibits some nugget effect, indicating that either depth is not accurately measured and/or that depth can change quite quickly with respect to the inter-sample distance.

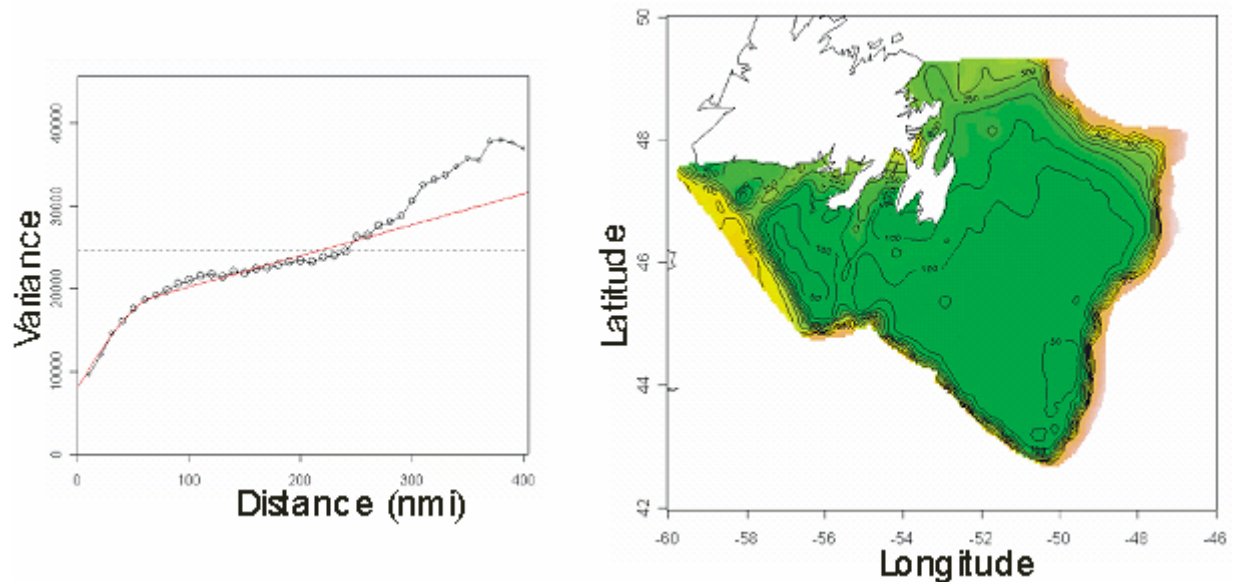


Fig. 25. Variogram (left), computed for all directions, of depths (m) sampled during the 1996-2002 Canadian spring bottom trawl surveys of the Grand Banks. The variogram model (red line) is superimposed on the experimental variogram. Kriging map of depth (with a unique neighborhood) shown at 50-m depth contour intervals (right).

4.6.3 Greenland halibut

Greenland halibut were distributed along the shelf edge during 1996-2002 (Fig. 26). This spatial characteristic must be taken into account in a spatial interpolation. Therefore, a transformation of the sampling coordinates, into a reference system with an axis across the shelf edge and another one along the shelf edge, was conducted (Fig. 26). Then, a complete geostatistical analysis including the variogram and the kriging was conducted using the new reference system. However, kriging has been back-transformed into a geographical reference system (Fig. 27).

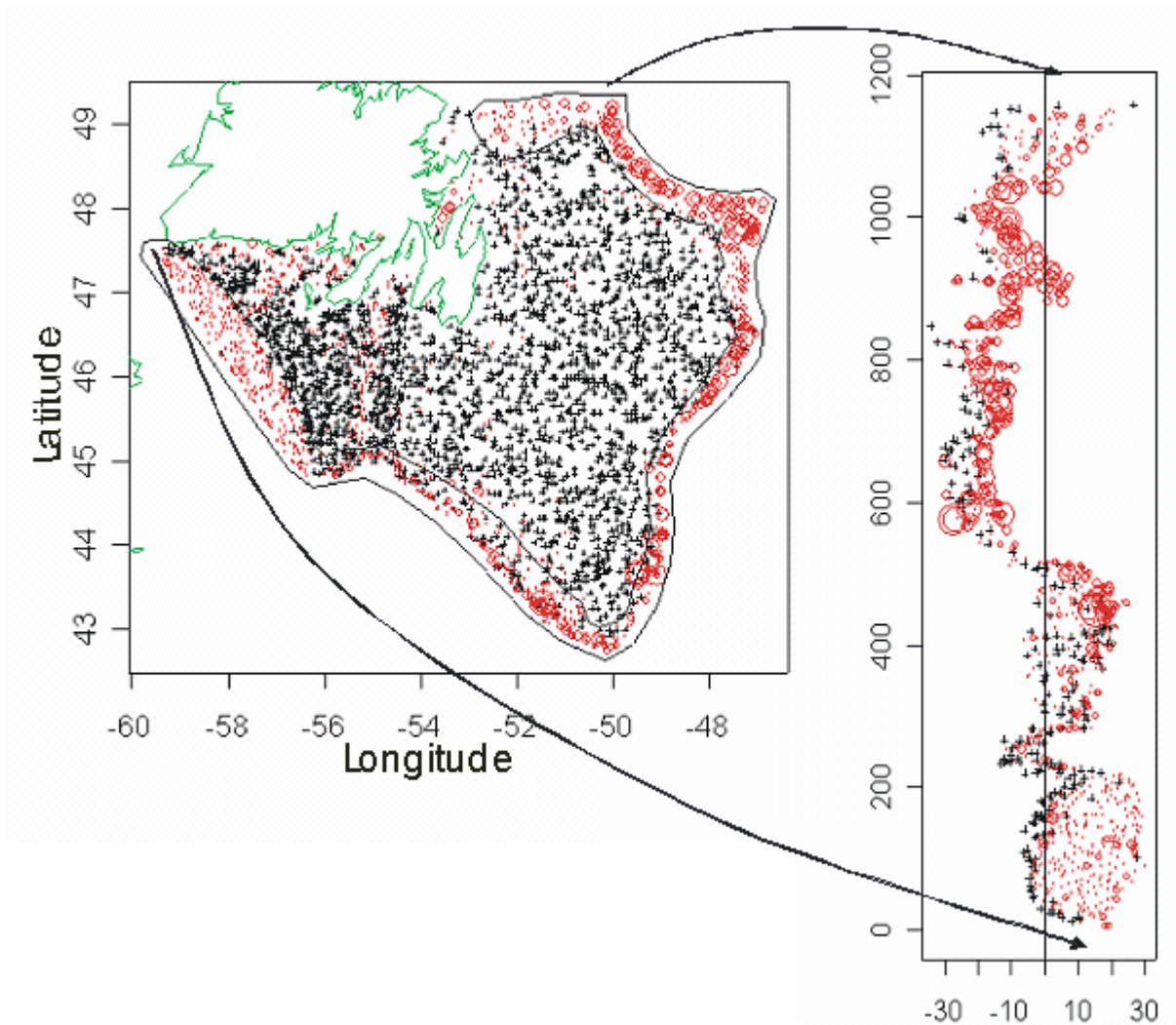


Fig. 26. Spatial distribution of Greenland halibut during spring of 1996-2002 based on Canadian bottom trawl surveys. Red circles increase in size with increasing halibut density and black crosses indicate stations with no halibut catch. The polygon indicates the selected area of occupancy (left) and the location of the selected points within this area are shown across (x axis, in nmi.) the shelf edge and along (y axis, in nmi.) the shelf edge (right).

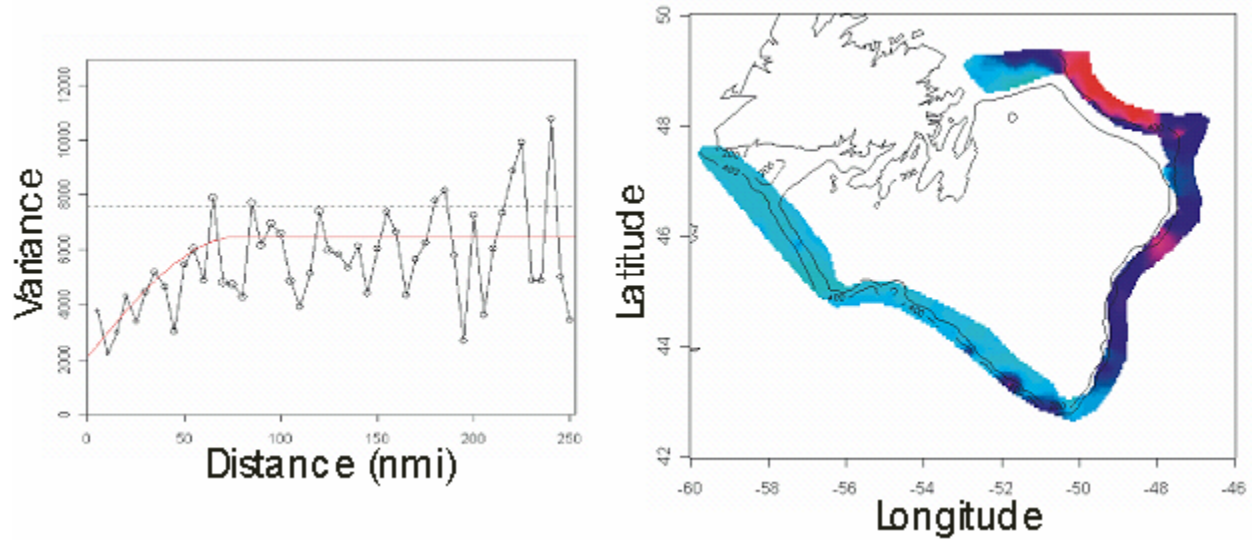


Fig. 27. Variogram (left) of Greenland halibut density along the shelf edge of the Grand Banks during spring of 1996-2002, based on Canadian bottom trawl survey data. The variogram model (red line) is superimposed on the experimental variogram. Kriging map (with a unique neighborhood) of Greenland halibut density in relation to the 200 m and 400 m isobaths (right).

4.6.4 CPUE

The potential use of CPUE (catch per unit effort) data to describe the spatial structure of yellowtail flounder was also assessed. Given the very dense concentration of the data points in space and time, a short distance lag was required (0.5 nmi.). To avoid the comparison between CPUE data collected too far apart in time, a time lag was also used. The level of local heterogeneity is still high (the nugget effect accounts for half of the variability) and a spatial structure of several nautical miles is apparent (Fig. 28).

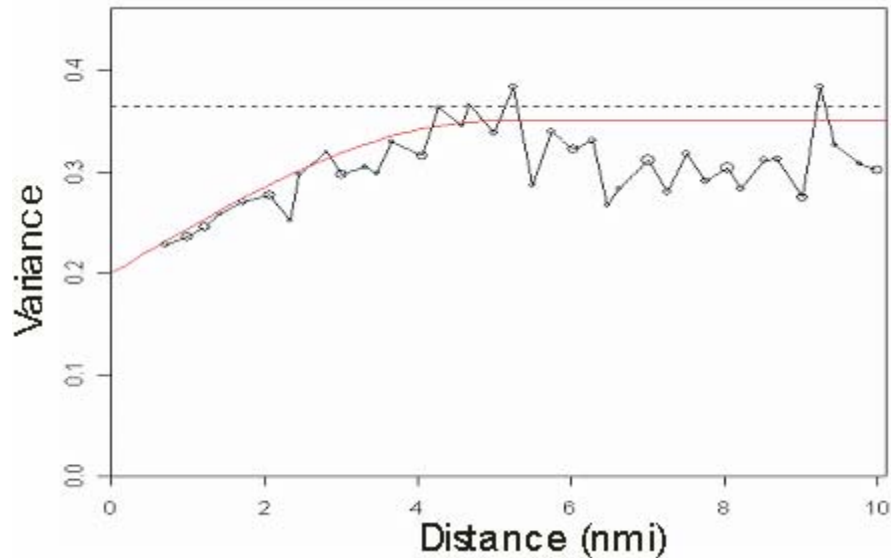


Fig. 28. Variogram of yellowtail flounder weekly CPUE on the Grand Banks during 2000, based on a distance lag of 0.5 nmi. The variogram model (red line) is superimposed on the experimental variogram.

5.0 CONCLUSIONS

The impetus for the workshop was an unfamiliarity of many Scientific Council members with a geostatistical presentation, GAM and kriging analyses of yellowtail flounder distribution on the Grand Banks (Walsh *et al.* 2001). As a result of this workshop, the participants now have a better understanding basic geostatistical concepts and methods, particularly kriging. In addition, the workshop provided members with freeware software tools and hands-on exercises that can be applied to geo-referenced data from the NAFO region in the future.

Participants discussed how they might apply the knowledge gained at the workshop. A recommendation was made to investigate efficient ways of incorporating mapping and geostatistical analyses into NAFO stock assessments, possibly by accessing these tools via an internet site, such as the GMBIS website, which already provides a mechanism for mapping user-defined geo-referenced data. However, it was also noted that the time commitment and programming knowledge required for this task represent potential obstacles to implementation. Multiple participants expressed positive feedback about the workshop, particularly with respect to covering such a complex topic, in a short time span, in an understandable way and with the use of NAFO data sets. Participants also noted that they felt that the knowledge gained at the workshop would now allow them to apply geostatistical analyses to their NAFO stock assessments.

6.0 REFERENCES

- ANON. 1999. Spans 7.1 Manual. Prospector TYDAC Research Inc. All rights reserved. Ed: Diane Burke. 108 p.
- BERTIN, J. 1981. Graphics and Graphic Information-Processing. Walter de Gruyter & Co. N.Y., 274 p.
- BETZ, D. 1988. Some Assembly Required - Embedded Languages. BYTE Nov. 1988, Vol **13**, No. 12, pg 409-416.
- BEZ, N. AND RIVOIRARD J. 2000. MS Indices of collocation between populations. In : Checkley, D.M., J.R. Hunter, L. Motos, and C.D. van der Lingen (eds). Report of a workshop on the use of Continuous Underway Fish Egg Sampler (CUFES) for mapping spawning habitat of pelagic fish. GLOBEC Report 14, 1-65 pp.
- BEZ, N. AND RIVOIRARD J. 2001. Transitive geostatistics to characterise spatial aggregations with diffuse limits: an application on mackerel ichthyoplankton. *Fisheries Research* **50** 41-58 p.
- BLACK, G.A.P. 2002. MS ACON Data Visualization Software User Manual - Version 9.15 - March 7, 2002. <http://www.mar.dfo-mpo.gc.ca/science/accon/>
- CONAN G. 1985. MS Assessment of shellfish stocks by geostatistical techniques. ICES CM 1985/K:30.
- GOHIN, F. 1985. MS Geostatistics applied to fish distributions as derived from acoustic surveys. Working Document, Acoustic Working Group Meeting, Tromsø, 22-24 May 1985.
- ICES, 1993. MS Report of the Workshop on the Applicability of Spatial Statistical Techniques to Acoustic Survey Data. 5 - 9 September 1991, Reykjavik, ICES Cooperative Report no 195.
- KULKA, D. W. 1998. MS SPANdex - SPANS geographic information system process manual for creation of biomass indices and distributions using potential mapping. DFO Atl. Fish. Res. Doc. 98/60. 28p.
- KULKA, D.W., AND PITCHER, D.A. 2001. MS Spatial and Temporal Patterns in Trawling Activity in the Canadian Atlantic and Pacific. ICES CM 2001/R:02. 57 p.
- LALOE, F. 1985. Contribution à l'étude de la variance d'estimateurs de biomasse de poissons obtenus par échantillonnage. *Océanogr. Trop.* **20** (2) : 161-169.
- LAUREC, A. 1977. Analyse et estimation des puissance de pêche. *J. Cons. Int. Explor. Mer*, **37**(2): 173-185.

- O'BRIEN, L. AND P. RAGO 1996. An application of the Generalized additive model to groundfish survey data with Atlantic cod off the northeast coast of the United States as an example. *NAFO Sci. Coun. Studies* **28**:79-95.
- RIVOIRARD, J., J. SIMMONDS, K.G. FOOTE, P. FERNANDES AND N. BEZ, 2000. *Geostatistics for estimating fish abundance*. Ed. Blackwell Science, Oxford, 206 p.
- WALSH, S. J., M. R. SIMPSON, J. MORGAN, K. S. DWYER AND D. STANSBURY 2001. MS Distribution of Juvenile Yellowtail Flounder, American Plaice and Atlantic Cod on the Southern Grand Bank: a Discussion of Nursery Areas and Marine Protected Areas. NAFO SCR Doc. 01/78 Serial No. N4457.
- SWARTZMAN, G.W., G.C. HUANG AND S. KALUZNY 1992. Spatial analysis of Bering Sea groundfish survey data using generalized additive models. *Can. J. Fish. Aquat. Sci.* **49**: 1366-1378.

7.0 ACKNOWLEDGEMENTS

The Scientific Council would like to thank the NAFO Contracting Parties for funding the geostatistics expert, Nicolas Bez, who served as the principal workshop instructor. The Council also thanks Canada and the EU for funding the travel costs of the other instructors; David Kulka, Jerry Black and Mark Simpson (Canada) and Reiner Schlitzer (EU, Germany). The quality of the workshop was greatly enhanced by the presence of these instructors. The Council is grateful to Canada for providing two information technology experts (Pete Rioux and Todd Janes) and a wireless LAN system that greatly improved the efficiency of the workshop. We deeply appreciated the assistance from Johanne Fischer (NAFO Executive Director) and the NAFO staff for their assistance with the organization of workshop logistics and hospitality.

APPENDIX 1. WORKSHOP AGENDA

DAY 1

0800-0900 Establish LAN connections and download software

0900-0930 Introduction

Workshop Objectives, Agenda and Introduction of Instructors (D. Kulka/L. Hendrickson)

0930-1000 Overview

Workshop overview and brief history of geostatistics use in fisheries stock assessment (N. Bez)

- Data visualization (mapping techniques)
- Interpolation techniques - Point to Surface Transformation (i.e. Contouring, Voronoi, Potential Mapping, Kriging)
- Overlay modeling
- Geostatistics

Data Visualization

1000-1015 Overview of mapping (D. Kulka)

The value of visualization of biological and environmental data in the marine context will be reviewed. Spatial data structure will be described and illustrated.

1015-1035 *Break*

1035-1200 Use of ACON software (G. Black)

Participants will use ACON software (freeware) to map and analyze survey data. Transformation and visualization of point patterns to surface distributions will be examined. Tessellation methods such as plotting Voronoi polygons and the use of Delaunay triangulation will be described.

1200-1330 *Lunch*

1330-1430 SPANS software demonstration (D. Kulka)

Key functions in SPANS (Spatial Analysis System), a GIS, will be demonstrated. Potential mapping, a point to surface transformation will be demonstrated. Potential mapping provides an optimal interpolated estimate for locations that were not sampled. The resulting classified surface (raster) facilitates analyses not possible with the original point data.

1430-1500 Generalized Additive Models (M. Simpson)

1500-1520 *Break*

1520-1830 Use of Ocean Data View software (R. Schlitzer)

Participants will use Ocean Data View software to map and analyze NAFO survey data. This exercise will involve exploration and visualization of oceanographic and other geo-referenced profile or sequence data.

DAY 2**Geostatistics****830-1030** Surface Overlay and Data Modelling (D. Kulka)

Demonstration of overlaying surfaces and modeling to examine spatial relationships or to calculate spatial statistics. Examples will include species co-occurrence, habitat preferences and biomass calculations using NAFO data.

1030-1050 *Break***1050-1200** Why should we use geostatistics ? (N. Bez)

- Sampling Theory (Cochran, 1977) is based on the assumption that the sample values can be modeled as independent and identically distributed random variables. Each of these concepts will be discussed. When this framework is not consistent with the characteristics of either the sampling or the data, one alternative method is to use geostatistics.
- Modeling and use the use of autocorrelation present in the sample values of a given variable. Kriging will be presented as a method to allow weighting of the data according to a) spatial structure, b) relative location in space, and c) position relative to the point or the polygon to be estimated.
- Analysis of spatially-correlated (multivariate geostatistics)

1200-1330 *Lunch***1330-1530** The variogram (N. Bez)

- Background on variance
- Decomposition of the variance into distance bins
- Random Functions
- Variogram definition
- Estimation of the variogram in practice
- Models
- Properties
- Interpretation of the spatial structure

1530- 1550 *Break***1550- 1730** Estimation and interpretation of experimental variograms (N. Bez)

- Exercises using R (freeware version of Splus) and geostatistical routines.

DAY 3**0830- 1030** The variogram, continued (N. Bez)

- The variogram as a tool to compute variances
- Nugget effect and the reduction of variance
- Cases where statistics is relevant (pure nugget effects)
- Weighted variograms

Kriging

- Principles
- Equations
- Kriging properties
- Illustrations
- Difference between the variable and its kriging

- Kriging weights
- Local estimation as opposed to global estimations

1030- 1050 ***Break***

1050-1200 Kriging exercises (N. Bez)

Objectives include modeling the variograms and using them for kriging. Changes in the output maps by varying the variogram parameters will be examined.

1200- 1330 ***Lunch***

1330-1530 Elements of the transitive geostatistical approach (N. Bez)

- Center of Gravity and Inertia to summarize survey maps and to describe a series of survey data.
- Elements of global estimation (estimation variances, survey design, etc.)
- Presentation of alternative software (EVA, Isatis)

1530-1550 ***Break***

1550-1700 Wrap-up discussion

APPENDIX 2. WORKSHOP PARTICIPANTS

Co-conveners:

Lisa Hendrickson
National Marine Fisheries Service, NEFSC
Woods Hole, MA 02543
USA
Phone: +508-495-2285
Fax: +508-495-2393
E-mail: lisahendrickson@noaa.gov

David W. Kulka
Science, Oceans & Envir. Br., Dept. of Fish. & Oceans,
166 Water St., P.O. Box 5667
St. John's, Newfoundland A1C 5X1
Canada
Phone: +709-772-2064
Fax: +709-772-5469
E-mail: kulkad@dfo-mpo.gc.ca

Canada

Black, J.A. Invert. Fish. Div., Mar. Fish Div., Fisheries & Oceans, P. O. Box 1006, Dartmouth, NS B2Y 4A2
Phone: +902 426-2950 – Fax: +902 426-1506 – E-mail: blackj@mar.dfo-mpo.gc.ca

Branton, R.M. Marine Fisheries Division, Fisheries & Oceans, P. O. Box 1006, Dartmouth, NS B2Y 4A2
Phone: +902 426-3537 – Fax: +902 426-1506 – E-mail: brantonb@mar.dfo-mpo.gc.ca

Bowering, W.R. Science, Oceans & Envir. Br., Dept. of Fish. & Oceans, P.O. Box 5667, St. John's, Nfld. A1C 5X1
Phone: +709-772-2054 – Fax: +709-772-4105 – E-mail: boweringr@dfo-mpo.gc.ca

Brodie, W.B. Science, Oceans & Envir. Br., Dept. of Fish. & Oceans, P.O. Box 5667, St. John's, Nfld. A1C 5X1
Phone: +709-772-3288 – Fax: +709-772-4105 – E-mail: brodieb@dfo-mpo.gc.ca

Dwyer, K.S. Science, Oceans & Envir. Br., Dept. of Fish. & Oceans, P. O. Box 5667, St. John's, Nfld. A1C 5X1
Phone: +709-772-05736 – Fax: +709-772-4188 – E-mail: dwyerk@dfo-mpo.gc.ca

Maddock Parsons, D. Science, Oceans & Envir. Br., Dept. of Fish. & Oceans, P. O. Box 5667, St. John's, Nfld. A1C 5X1
Phone: +709-772-2495 – Fax: +709-772-4188 – E-mail: parsonsd@dfo-mpo.gc.ca

Morgan, M. J. Science, Oceans & Envir. Br., Dept. of Fish. & Oceans, P. O. Box 5667, St. John's, Nfld. A1C 5X1
Phone: +709-772-2261 – Fax: +709-772-4105 – E-mail: morganj@dfo-mpo.gc.ca

Orr, D. C. Science, Oceans & Envir. Br., Dept. of Fish. & Oceans, P. O. Box 5667, St. John's, Nfld. A1C 5X1
Phone: +709-772-7343 – Fax: +709-772-4105 – E-mail: ord@dfo-mpo.gc.ca

Simpson, M.R. Aquatic Resour. Div., Dept. of Fish. & Oceans, P.O. Box 5667, St. John's, Nfld. A1C 5X1
Phone: +709-772-5314 – Fax: +709 772-5469 – E-mail: simpsonmr@dfo-mpo.gc.ca

Stansbury, D.E. Science, Oceans & Envir. Br., Dept. of Fish. & Oceans, P. O. Box 5667, St. John's, Nfld. A1C 5X1
Phone: +709-772-0559 – Fax: +709-772-4105 – E-mail: stansburyd@dfo-mpo.gc.ca

Walsh, S.J. Science, Oceans & Envir. Br., Dept. of Fish. & Oceans, P. O. Box 5667, St. John's, Nfld. A1C 5X1
Phone: +709 772-5478 – Fax: +709 772-4105 – E-mail: walshs@dfo-mpo.gc.ca

Estonia

Saat, T. Estonian Marine Institute, 18b Viljandi Road, EE-3600 Tallinn
Phone: +372 6267 402 – Fax: +372 6267 417 – E-mail: tsaat@sea.ee

Faroe Islands

Nicolajsen, A. Fiskorannsóknarstovan, Noatun, Postboks 3051, FR-1100, Torshavn
Phone: +298 345 3900 – Fax: +298 35 3901 – E-mail: arninic@frs.fo

France

Bez, N. Centre de Géostatistique, Ecole des Mines de Paris, 35 rue Saint Honoré, 77305 Fontainebleau
Phone: +33 1 6469 4956 Fax: +33 1 6469 4705 – E-mail: nicolas.bez@ensmp.fr

Germany

Schlitzer, R. Alfred Wegener Institute for Polar & Marine Research, Columbusstrasse, 27568 Bremerhaven
Phone: +49 471 4831 1559 – E-mail: rschlitzer@awi-bremerhaven.de

Stein, M. Institut für Seefischerei, Palmallee 9, D-22767 Hamburg
Phone: +49 40 38905 174 – Fax: +49 40 38905 263 – E-mail: stein.ish@bfa-fisch.de

Greenland

Wieland, K. Greenland Institute of Natural Resources, P. O. Box 570, DK-3900, Nuuk
Phone: +299 32 1095 - Fax: +299 32 5957 - E-mail: wieland@natur.gl

Portugal

Avila de Melo, A. Inst Instituto Nacional de Investigacao Agrária e das Pescas (INIAP/IPIMAR),
Av. de Brasília, 1449-006 Lisbon, Portugal
Phone: +351 21 302 7000 – Fax: +351 21 301 5948 – E-mail: amelo@ipimar.pt
Alpoim, R. Instituto Nacional de Investigacao Agrária e das Pescas (INIAP/IPIMAR),
Av. de Brasília, 1449-006 Lisbon, Portugal
Phone: +351 21 302 7000 – Fax: +351 21 301 5948 – E-mail: ralpoim@ipimar.pt

Russian Federation

Babayan, V.K. Russian Federal Research Institute of Fisheries & Oceanography (VNIRO), 17, V. Krasnoselskaya
Moscow, 107140
Phone: +70 95 264 6983 – Fax: +70 95 264 9183 – E-mail: vbabayan@vniro.ru
Gorchinsky, K.V. Knipovich Polar Research Institute of Marine Fisheries and Oceanography (PINRO),
6 Knipovich St., Murmansk 183763
Phone: +7 (8152) 47 2532 – Fax: +7 (8152) 47 3331 – E-mail: inter@pinro.ru
Kokovkin, L. Representative Russian Federation on Fisheries, 47 Oceanview Dr., Bedford, N.S., Canada B4A 4C4
Phone: +902 832-9225 – Fax: +902 832-9608 – E-mail: rusfish@ns.sympatico.ca
Shibanov, V.N. Knipovich Polar Research Institute of Marine Fisheries and Oceanography (PINRO),
6 Knipovich Street, Murmansk 183763
Phone: +7 8152 47 26 14 – Fax: +7 8152 47 33 31 – E-mail: inter@pinro.murmansk.ru
Sigaev, I.K. Atlantic Scientific Research Institute of Marine Fisheries and Oceanography (AtlantNIRO)
5 Dmitry Donskoy Street, Kaliningrad 23600
Phone: +70 112 22 5547 – Fax: +70 112 21 9997 – E-mail: west@atlant.balmet.ru

Spain

Gonzalez-Costas, F. Instituto Español de Oceanografía, Aptdo 1552, E-36280 Vigo (Pontevedra), Spain
Phone: +34 9 86 49 2111 – Fax: +34 9 86 49 2351 – E-mail: fernando.gonzalez@vi.ieo.es
Vazquez, A. Instituto de Investigaciones Marinas, Eduardo Cabello 6, 36208 Vigo
Phone: +34 9 86 23 1930 – Fax: +34 9 86 29 2762 – E-mail: avazquez@iimcsic.es

USA

Mayo, R.K. National Marine Fisheries Service, NEFSC, 166 Water St., Woods Hole, MA 02543
Phone: +508-495-2310 - Fax: +508-495-2393 - E-mail: ralph.mayo@noaa.gov
Serchuk, F.M. National Marine Fisheries Service, NEFSC, 166 Water St., Woods Hole, MA 02543
Phone: +508 495 2245 – Fax: +508 495 2258 – E-mail: fred.serchuk@noaa.gov

Note: Instructors names are shown in Bold



Participants of Workshop on "Mapping and Geostatistical Methods for Fisheries Stock Assessment, 10-12 September 2003.

Left to Right: Dave Kulka, Bill Brodie, Kai Wieland, Cindy Kerr, Steve Walsh, Don Stansbury, Mark Simpson, Karen Dwyer, Dawn Maddock Parsons, Ralph Mayo, Jerry Black, Dave Orr, Lisa Hendrickson, Ricardo Alpoim, Fred Serchuk, Antonio Avila de Melo, Dorothy Auby, Reiner Schlitzer, Joanne Morgan, Nicolas Bez, Bob Branton, Toomas Saat, Arni Nicolajsen, Ray Bowering, Antonio Vazquez, Manfred Stein, Igor Sigaev, Vladimir Shibanov, Konstantin Gorchinsky, Vladimir Babayan.

Front Centre: Fernando Gonzalez-Costas

Missing: Tissa Amaratunga, Leonid Kokovkin



Organizers and Instructors:

Mark Simpson, David Kulka, Jerry Black, Nicolas Bez, Reiner Schlitzer and Lisa Hendrickson.