NAFO Sci. Coun. Studies, 28: 55–63

Overview of Environmental Aspects for the Analysis of Fish Stocks Using Bottom Trawl Surveys

Manfred Stein
Bundesforschungsanstalt für Fischerei
Institut für Seefischerei, Palmaille 9, 22767 Hamburg, Germany

Introduction

Bottom trawl surveys

Oceanography, if at all included in bottom trawl surveys, tries to map the environmental background in relation to the surveyed species. During such surveys, often before or after a haul, CTDs or Nansen bottles are used to sample the vertical structure of the thermo-haline fields of the water column. If equipped with more sophisticated instrumentation, profiles of other water mass properties are sampled, e.g. oxygen, sound-velocity, attenuation, current velocity. Analysis of the raw data is achieved using standard calibration routines for CTD or Nansen bottle data.

Unfortunately, the clean oceanographic data obtained during bottom trawl surveys are very often not analysed in a way to fulfill the initial demands, i.e. map the environmental background of the surveyed species. The reasons for this are manyfold: lack of time, lack of manpower, lack of skill, or lack of requirement (due to other priorities). Additionally there are difficulties in how to analyse oceanographic data which are sampled at random stations, since normally oceanographic data are sampled along oceanographic sections or at standard stations and are analysed according to classical routines. But there are valuable data collected in trawl surveys, and the task of the present is to consider how to use them. In the following, three major subjects items will be covered:

- Data handling by Data Information Systems
- Data analysis by Interpolation software
- Some theoretical background information

Let us assume that around 100 CTD stations have been made during a bottom trawl survey, and these data are now in a ready format (Institute format, GF3 format, manufacturers format).

Subject Items

Data Handling by Data Information Systems

Different kinds of software, ranging from home-made to professional software, are in use to extract information from this data set. In my Institute I have developed a software, called ODISys, which is an Oceanographic Data Information System. This software is designed to take the user step by step.

The user is first asked to key in the type of files where to search for information (e.g. **.STD** files: these files contain information on oceanographic profiles at oceanographic standard depths: 0, 10, 20, 30, 50, ... , 200 m and so on). The program is then able to search all files with the extension **.STD**, for the desired information. The next entry is to define the area delimited by the northern, southern, eastern and western boundaries. The water depth for the search has to be given along with a potential upper and lower limit for the depth range (e.g. 200 m plus/minus 10 m, if data are not on standard depth levels). Finally, the parameter which is to be extracted from the data base has to be given.

The software continues with a display of the coastal boundaries of the area under investigation. After some search time (the time depends on the hardware and the amount of data profiles) station positions and further information is displayed on the screen giving the min/max of the retrieved parameter, the period of time (year, month), and the name of the file in which the retrieved data are stored.

It is this file which forms the basis for the subsequent analyses (e.g. by spreadsheet or by SURFER). It is an ASCII file (Table 1) with variables delimited by blanks.

In Table 1 the first line gives the area from which data are retrieved (i.e. the N, S, E, W boundary), depth of retrieved data (50 m), retrieved parameter (1: measured temperature). The subsequent lines give sample location (longitude, latitude), temperature, year (e.g. 95: is 1995), month (10 is October) of the retrieved temperature data at 50 m depth (this example shows only 6 records).

TABLE 1: Example of data retrieved from WH164.

500000N	700000N	0440000W	0700000W	50.00 1
-45.33	59.00	6.32	95	10
-44.77	59.27	5.07	95	10
-44.51	59.45	5.41	95	10
-44.15	59.63	2.97	95	10
-44.23	59.67	2.39	95	10
-44.62	59.71	1.41	95	10
.....	.....	....	..	..

Note: When working under SURFER with this file, the first line (500000N 700000N 0200000W 0700000W 50.00 1) has to be scratched.

Data analysis by Interpolation software

For the purpose of this presentation, I will first deal with a data set which was obtained along a regular grid (box N/S station distance 15 naut. miles), and secondly I will deal with a data set derived during the 1995 bottom trawl survey off West Greenland. To avoid confusion, the example Figures have been trimmed to show the most necessary information (e.g. no text, no axis labelling). I promise you, the actual outputs are nice looking figures.

Figure 1 shows a grid of 37 stations sampled in an area of 2° latitude and 4.5° longitude. The retrieved parameter was salinity, and the observation points (locations) are indicated by a cross symbol.

Automatic contouring has two distinct problems of gridding and plotting: Gridding is required whenever data points are not on a regular mesh (our first data example, although sampled along a fixed grid, is in reality not on a regular mesh, due to position inaccuracy). However, even when data are sampled on a regular grid, a smaller mesh might be necessary. Plotting is finding points along lines of constant parameter-values (e.g. salinity in our first example). Plotting includes smoothing and labeling.

The most critical step is gridding. The task is to find the optimal interpolation procedure to assess the goodness-of-fit.

Figure 2 gives an example of a grid which is applied by the SURFER software as standard (i.e. no change of Grid Size). The grid is 25 by 25, thus giving 625 node points at which interpolation has to be achieved. As concerns regional coverage, one can easily see where the deficiencies of our data set lie: the southeastern corner of the box.

At this point it seems suitable to give some theoretical background information:

Interpolation procedures use linear combination of the data. If we assume sample points in the *p*-dimensional space as x_i, known values as $z_i = z(x_i)$, then a linear combination would be kind of a summation

$$z = \sum_i \lambda^i z_i \tag{1}$$

λ^i are coefficients, or weights, for the measured values z_i. The least-square estimator is of type (1). Distance Weighting Functions are a sensible, easy scheme for weighting data points to ascribe higher weights to closer points. The choice of distance weighting function is more or less a matter of personal preference, of tradition, or of confidence in the advice of "influential authorities" (Delfiner *et al.*, 1975). Distance weighting functions are impaired by a major insufficiency: they fail to discriminate redundant information.

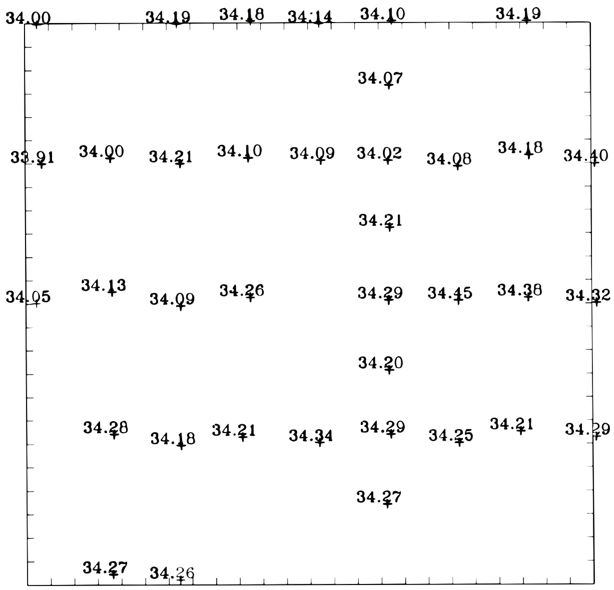

Figure 1.

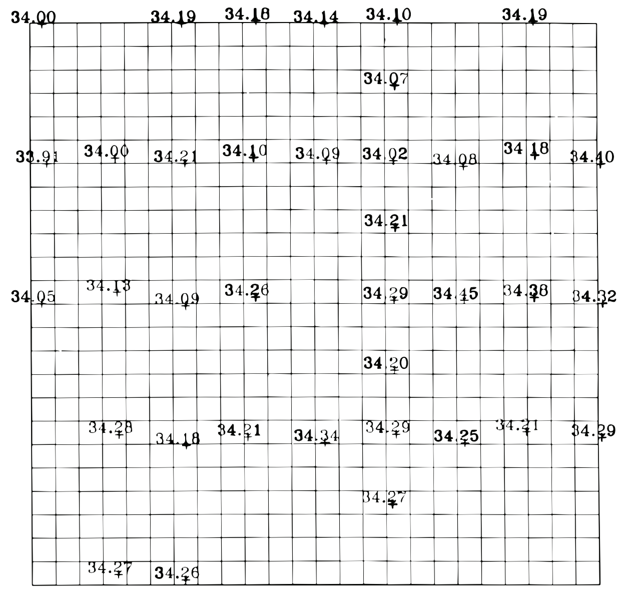

Figure 2.

Let us take an example: three equidistant points (A, B, C) from the estimated node point 0. If the variable is isotropic and homogeneous (without trend), any procedure will ascribe the same weight (1/3) to A, B, C, provided the weights add up to one. There is, however, a problem with clusters.

Conclusion: it is necessary to take into account not only the situation of sample points with respect to the estimated (node) point, but also interrelationships between sample points themselves. The least-squares procedure achieves this. It disregards, however, the structure of the variable under study.

If a variable is very continous, e.g. 500 millibar pressure surface, the fitted surface should be able to honor the experimental points. However, there are irregular variables (mineral grades in mines, bathymetric depths, or local temperature in a layer of the ocean). When data show such considerable scatter even at a small scale, it makes no sense to constrain the surface to pass through the experimental points. Some type of filtering is required, but which filtering method would be best? It is most important to know precisely and beforehand what is to be filtered off and what is to be kept. The least-square technique yields "trends" which are sometimes hard to interpret. As emphasized by Matheron (1971) "the impression is often created that this famous trend is nothing more than a numerical result brought about by the mode of operation – i.e. perhaps a pure and simple artefact."

This points at some peculiar problems which often arise with SURFER products. However, it is my hope that this Workshop will help you to avoid most of these problems in future.

The optimum interpolation method is Kriging (Upton *et al.*, 1985).

The approach generally referred to as Kriging (on account of the work of the geostatistician D. G. Krige, 1966; see also Krige *et al.*, 1969) aims to produce accurate predictions by interpolating from localities with known z_i -values to 'prediction localities' for the missing values. Many applications of these methods have been in the mining industry, particularly South African gold mines, where accurate prediction is essential and where a smooth function such as a trend surface is an inadequate representation of the variation typically associated with, say, an ore-body. The Z-surface formed by ore-grade variations around a mine tends to be continuous and intensely locally variable, so that trend surfaces *per se* provide predictions that are only broadly accurate, since the estimates z are the consequence of fitting the surface to the entire data set and assuming normal errors. Although an ore-body or similar surface may actually contain such a large-scale component, a more effective prediction technique is to combine knowledge of the trend at the prediction locality (z) with a local component (l_i) derived from the covariance structure of the residuals from that trend.

By definition, the trend will describe the large-scale fluctuation of the surface and the local or residual component will represent small-scale fluctuation with a tendency to peter out.

I will not go into further details on the mathematics of the Kriging equations. For those who are interested, I have listed in the references some literature which might help to understand why Kriging provides the best linear unbiased estimate built on the observed data.

Instead, let us continue with our data example: Figure 3 gives a first impression of plotting after gridding has been achieved by the quick but raw Inverse Distance Method. No changes are inferred to the settings of the SURFER program, except for the font which has been chosen as Times Roman.

Figure 4 has been prepared with standard settings, however with the Kriging Method. The isolines are much smoother than in Fig. 4, and the impression we receive is a "combination of knowledge of the trend at the prediction locality (z) with a local component". The Kriging option does not infer so much singular, closed areas, which as stated above are artefacts.

Application of a Spline routine on the "Kriged Grid" is shown in Fig. 5 (note the finer grid in this application of 73 by 73).

As a last step of plotting – before adding land contours and text – Fig. 6 yields equidistant isolines. A label option (Labeled contour line frequency) of 1 is chosen, i.e. every contour line is labeled.

The purpose of my lecture is: The Environmental Aspects of the Analysis of Fish Stocks Using Bottom Trawl Surveys. So far, we have dealt with one side of the coin, the technical aspect.

The oceanographic analysis of the model data set as displayed in the previous figures enables, however, to get a closer view of the distribution of water masses, fronts, currents and so on. Standard depth level 50 m (Fig. 7) indicates a clear discrimination between low saline water in the northwestern corner of the grid, and distinctly higher salinities in the southeastern corner. These two domains are separated by a meandering front, a confluence zone. This confluence also emerges as a dynamic oceanographic feature. From geopotential anomaly of the CTD dataset, the 0/300 value was taken to give an estimate for the surface currents referenced to the 300 dbar pressure level (Fig. 8). The flow direction is generally from west to northeast (meander), with a return flow in the northwestern corner, and low baroclinic activity in the southern part of the grid. The isolines are given in dynamic metres.

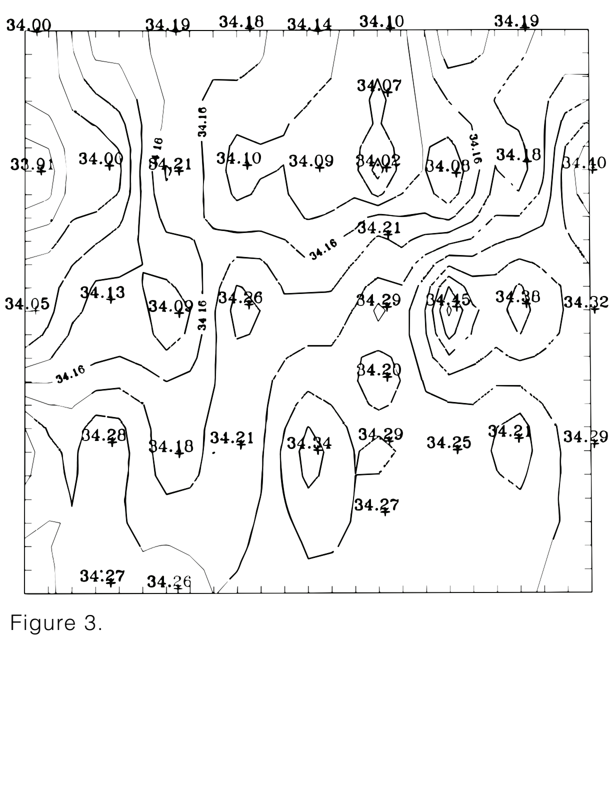

Figure 3.

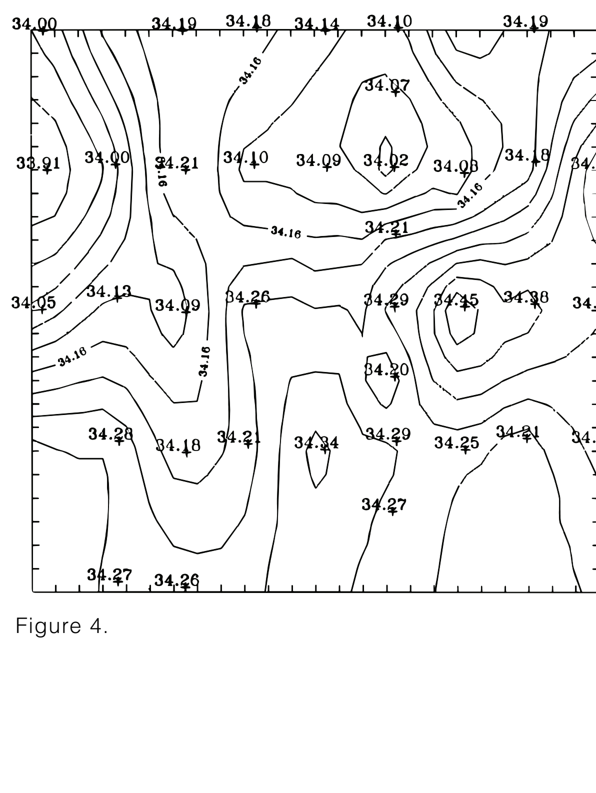

Figure 4.

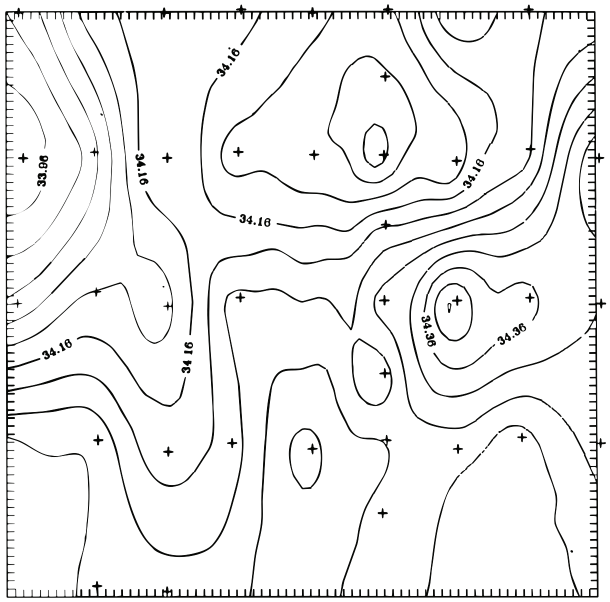

Figure 5.

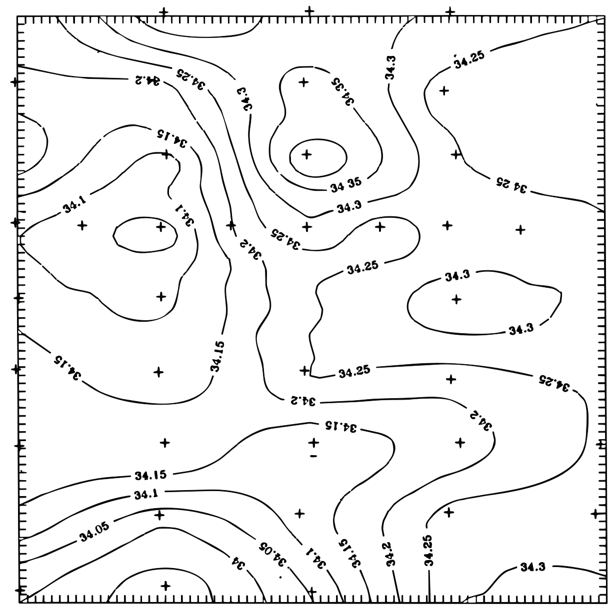

Figure 6.

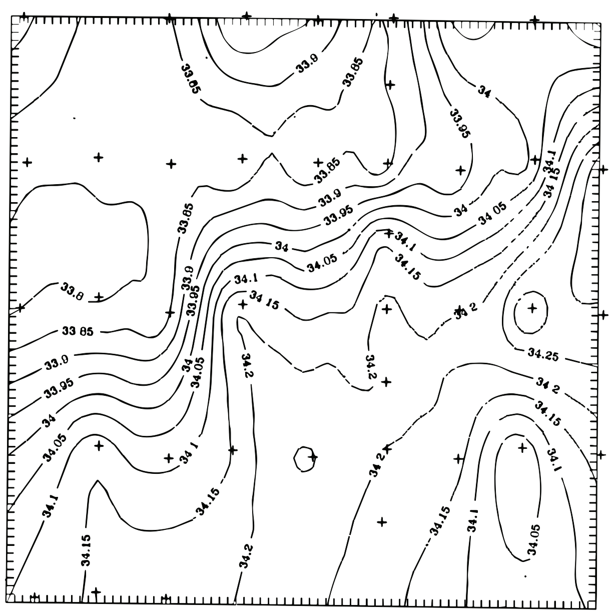

Figure 7.

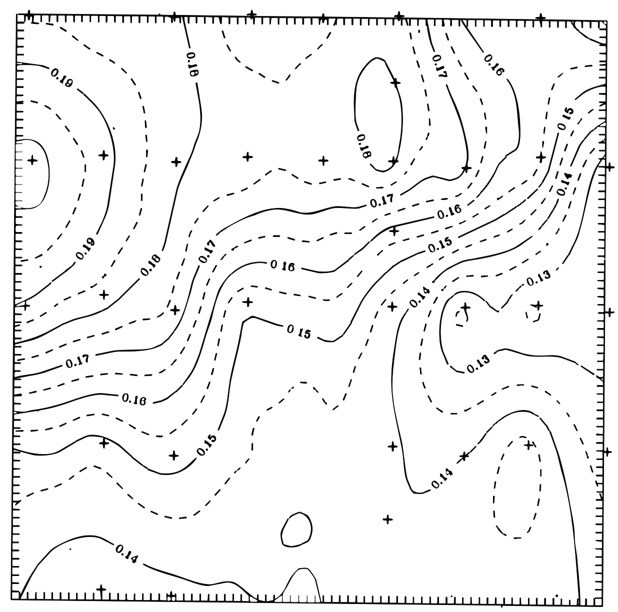

Figure 8.

The present data set originates from the Antarctic Ocean. It is part of my Institutes' oceanographic data base. I have chosen this data set since it can be used to demonstrate:

- oceanographic features in the vicinity of an island shelf (topographic steering of a water mass boundary, the Weddell–Scotia Confluence)
- open ocean boundaries (this is in contrast to the e.g. Greenland oceanographic data set)
- availability of a near-regular observation grid on a rectangular basis
- availability of parameter set (T, S, Sv, Sigma, SDD)

- impact on distribution of biota, e.g. planctonic larvae
- clear cut between different oceanographic regimes

Our second example originates from CTD data sampled during the bottom trawl survey in autumn 1995 off West Greenland (Table 1). Forty-six stations were sampled off West Greenland, with temperature ranging from 1.21 C to 6.32 C at 50 m (Fig. 9). In contrast to our first example, the second data set follows the continental margins off the West Greenland banks, and thus the data points are grouped along one diagonal of the SURFER rectangle.

Let us see how SURFER deals with this irregular spaced data set. The first approach with the Inverse Distance Method yields no convincing results (Fig. 10). The structure of the West Greenland Current system does not at all emerge from this picture. Let us try the Kriging Method. As can be seen from Fig. 11, a filamentous structure is outlined along the data grid. At this stage of plotting we will not care about all the fancy contouring outside the southeast-northwest diagonal. The Spline method applied to the Kriged data in Fig. 11 yields a filamentous structure along the southwestern Greenland coast (Fig. 12).

The task we now have is to get rid of the redundant part of the isolines. If we assume that within the area covered by the observed data points, Kriging provides the best linear unbiased estimate built on these observed data, then Blanking of the redundant part of the figure is suitable. SURFER provides a method to do this task in a simple way (Fig. 13). Delineated by the user defined boundary positions around the data grid, SURFER blanks with individually created **.BLN** files, the redundant information outside of this user defined area.

Adding to Fig. 13 the appropriate coastal contour file gives Fig. 14. Note that the land contours of southwest Greenland do not fit exactly to the frame (the coastal file is bounded by lat/long 59N–63N, 44W–53W, while the plotted area is somewhat smaller). To achieve the coastal contour file to fit the gridded area, a little trick is required. So far, we did not take into consideration the Limits function during the Gridding procedure. If the Limits are set to the actual limits of the coastal contour file, the plot looks very much acceptable, Fig. 15, at least to me!

Text, and labels have been added (Figs. 15, 16), and a special Search method is applied to the data set in Fig. 16 to reveal the filamentous character of the region even more clearly.

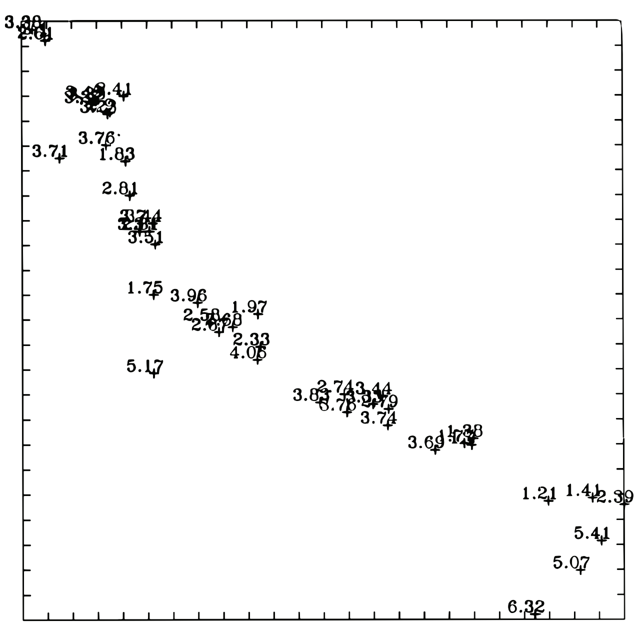

Figure 9.

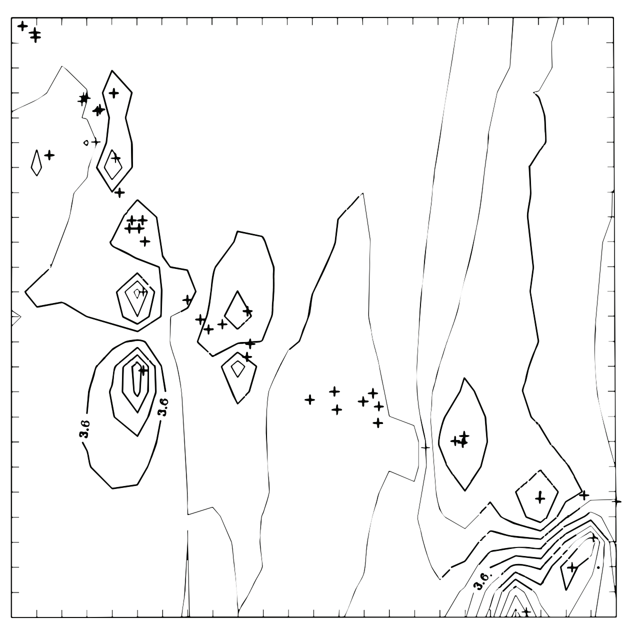

Figure 10.

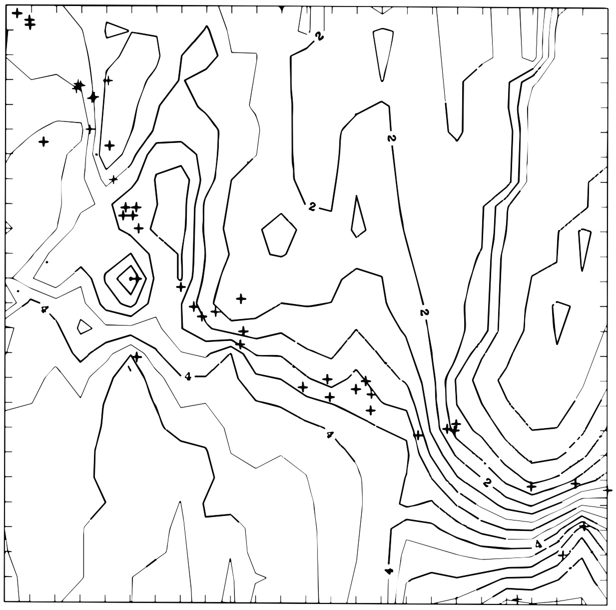

Figure 11.

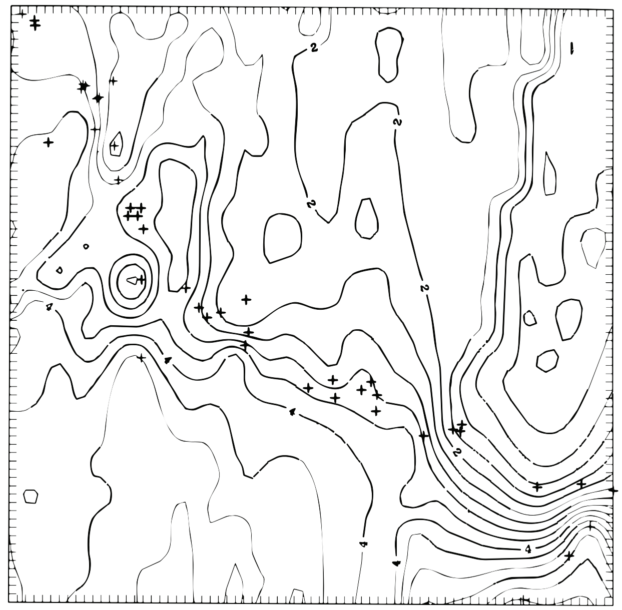

Figure 12.

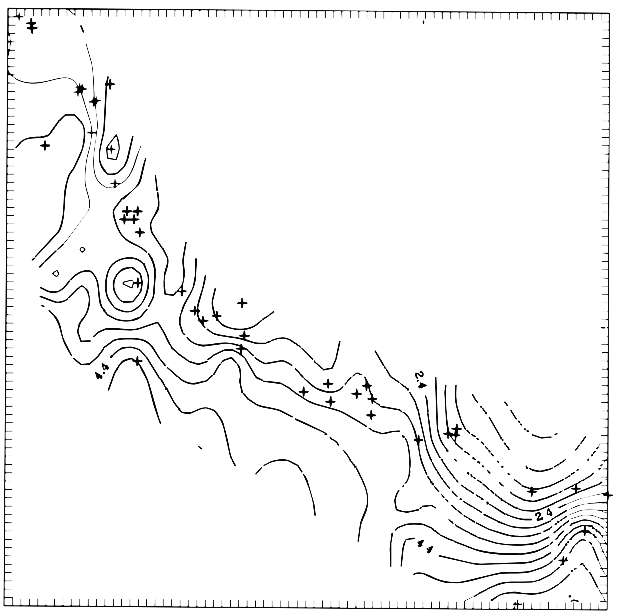

Figure 13.

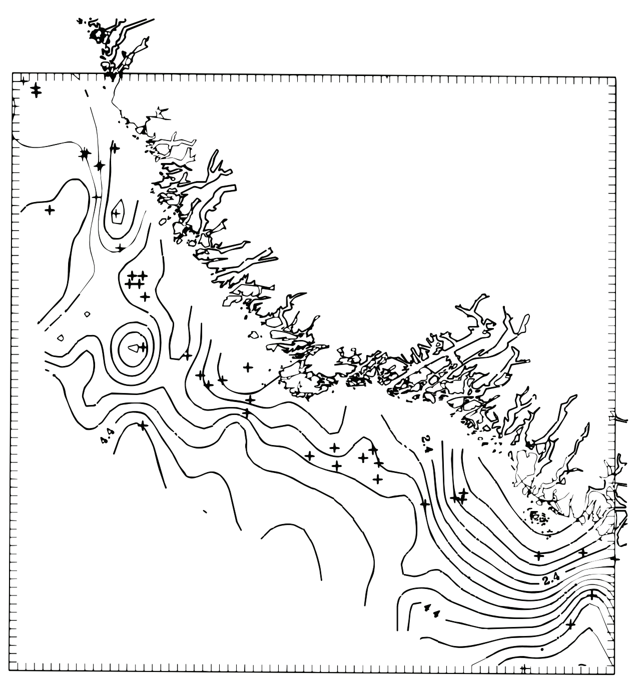

Figure 14.

For the bottom water layer, a data set has been extracted from the oceanographic data base by means of ODISys, and treated in the same way as the previous data set from West Greenland at 50 m depth. Data from the deep section stations are omitted.

Both (Figs. 17 and 18) indicate that the bottom water layer is dominated by different regimes: cold, low saline on the banks and in the vicinity of 61°N, 50°W, and warm, saline at the shelf break. Whereas the banks are covered by Polar and mixed water, the off-slope area is under the influence of the warm and saline Irminger component of the West Greenland Current system.

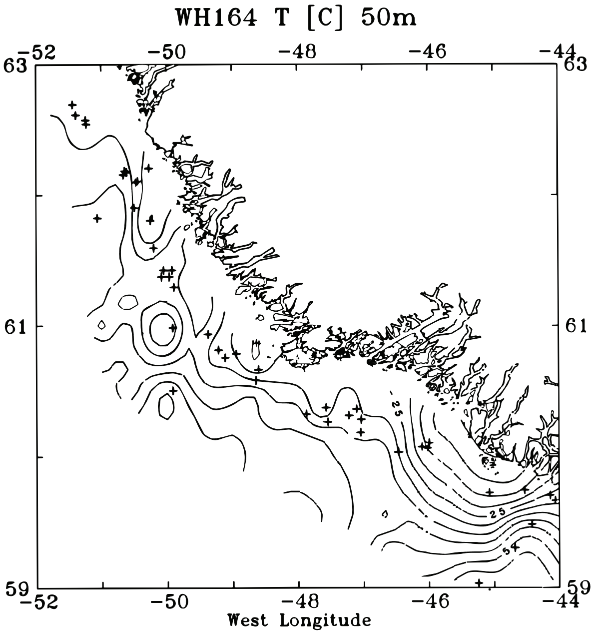

Figure 15.

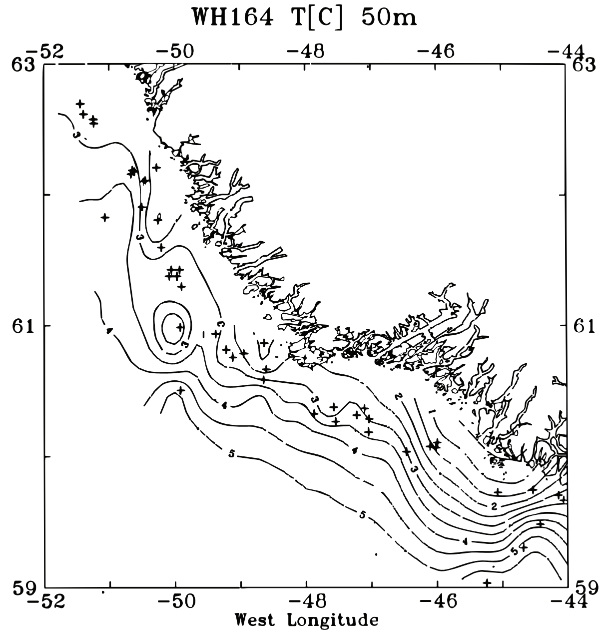

Figure 16.

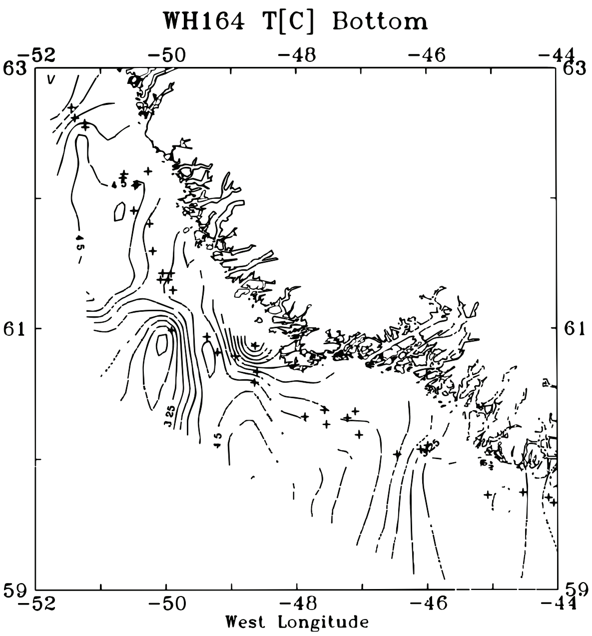

Figure 17.

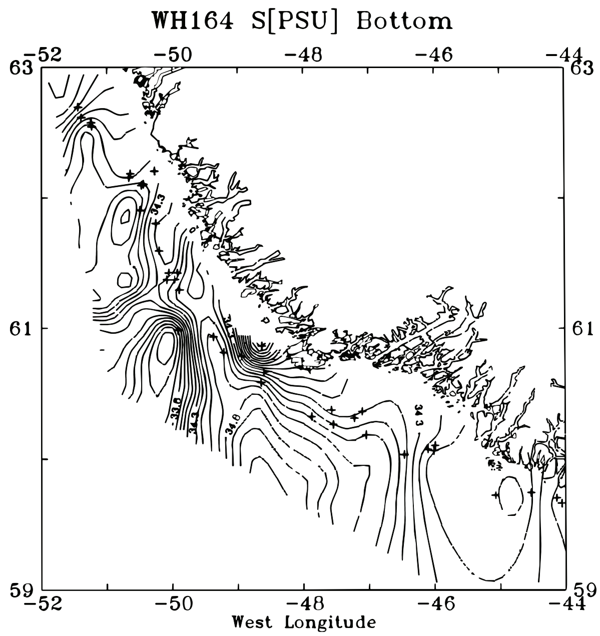

Figure 18.

It should be of interest to plot the species distribution of the demersal fish on top of these graphs to see whether there is any coincidence in the distribution of environmental parameters and the biota.

References

DELFINER, P., and J. P. DELHOMME. 1975. Optimum Interpolation by Kriging, *In* Display and Analysis of Spatial Data, J. C. Davis, and M. J. McCullagh (Eds.). Wiley: Chichester, 96–114.

KRIGE, D.G. 1966. Two-dimensional weighted moving average trend surfaces for ore evaluation, *In* Proceedings of the Symposium on Mathematical Statistics and Computer Applications in Ore Valuation, Johannesburg, 13–38.

KRIGE, D.G., M. I. WATSON, W. J. OBERHOLZER, and S. R. DuTOIT. 1969. The use of contour surfaces as predictive models for ore values, *In* A Decade of Digital Computing in the Mineral Industry, A. Weiss, (Ed.). American Institute of Mining, Metallurgical and Petroleum Engineers: New York, 127–161.

UPTON, G. J. G., and B. FINGLETON. 1985. Prediction methods related to Kriging, *In* Spatial Data Analysis by Example, Volume I Point Pattern and Quantitative Data, John Wiley & Sons: Chichester, 366–371.