Northwest Atlantic Fisheries Organization

Serial No. N1068 NAFO SCR Doc. 85/93

SCIENTIFIC COUNCIL MEETING - SEPTEMBER 1985

A Comparison of Survey Stratification Schemes Based on Depth and on Historical Spatial Dispersion

by

S. Gavaris

Department of Fisheries and Oceans, Marine Fish Division
Fisheries Research Branch, Biological Station
St. Andrews, New Brunswick, Canada, E0G 2X0

and

S. J. Smith

Department of Fisheries and Oceans, Marine Fish Division
Fisheries Research Branch, Bedford Institute of Oceanography
P. O. Box 1006, Dartmouth, Nova Scotia, Canada B2Y 4A2

Introduction

By the late 1960's the value of survey data for management purposes was recognized. Among the uses of survey data Grosslein (1969a) listed the following as principal objectives:

1. To monitor fluctuations in structure and size of fish populations.
2. To assess the fish production potential of Atlantic coastal waters.
3. To determine environmental factors controlling fish distribution and abundance.
4. To provide basic ecological data on fishes necessary to understand interrelationships between fish and their environment.

As one of the first steps towards establishing standardized and coordinated ICNAF-wide surveys, an ad hoc working group of STACRES endorsed the stratified-random design (Redbook, 1970). The advantages of the stratified-random design were reviewed by Grosslein (1969b) and more recently by Doubleday (1981). Foremost amongst the reasons for preferring the stratified-random design were:

1. Produces unbiased estimates of abundance indices with associated estimates of precision.
2. Sampling can be spread out over the area of study.
3. Sampling rates can be varied between strata to improve precision.
4. Strata can be aggregated to form domains of study.

SPECIAL SESSION ON BIOLOGICAL SURVEYS

The history of trawling surveys in Divisions 4VWX was reviewed by Halliday and Koeller (1981). During 1969 and 1970 the groundfish survey program was modified in accordance with the ICNAF recommendations. Recognizing the large scale trends in groundfish abundance with respect to hydrographic and bathymetric conditions, a stratification scheme based on depth was developed for the Scotian Shelf (Fig. 1). The allocation of stations to strata follows a prescribed schedule (Table 1).

Recent assessments of cod in Divisions 4VsW have relied extensively on abundance estimates from surveys (Sinclair and Gavaris, 1985; Gagné et al., 1984, 1983).

In this study, past data on the distribution and density of cod in 4VsW were used to define an alternative stratified-random design. We focussed on a single objective, the estimation of abundance indices for adult cod (ages 5-12) and compared the relative efficiency of the present and alternative sampling design with respect to simple random sampling. The primary domain considered was 4VsW which is the stock area. In addition, results for 4Vs and 4W, were examined separately to investigate the possibility of differing patterns in the two areas.

Alternative Sampling Design

Construction of Strata

The best characteristic for the construction of strata is the frequency distribution of the variable of interest (Cochran, 1977, p. 127). In our case the frequency distribution of the number of age 5+ - 12 cod caught is unknown. The next best characteristic is the frequency distribution of a quantity which is highly correlated with the variable of interest. The present stratification scheme used depth as such a quantity. In this study the use of historical density patterns to estimate a frequency distribution is explored, based on the premise that relative density from one year to the next was related. Scott and Gavaris (1985) concluded that, though distribution patterns vary from year to year, persistent features exist.

It was recognized that the observations from a single year would not be sufficient to characterize a frequency distribution or define geographical areas of low and high abundance. To use several years concurrently, the data from each survey was normalized by dividing each observation by the mean for that survey. In this way the expected contribution of each observation with respect to relative density would be equal. The data were examined in 5-year periods, 1970-74, 1975-79 and 1980-84, as discussed in Scott and Gavaris (1985).

The method described in Cochran (1977, p. 129) was used to define the limits of the normalized density values for strata. This involves the tabulation of the cumulative square root of the frequency and partitioning it into segments of equal range. Several large values of normalized density were excluded in order to facilitate analysis. Their exclusion has minimal effect on the outcome. The cumulative square root frequencies in each of the 5-year periods showed similar patterns (Table 2). Following Cochran's (1977, p. 142) suggestion that little reduction in variance could be expected by defining more than six strata, the frequencies were partitioned into six segments, however, the first two were not separable without finer classification of the frequencies. The resultant limits of normalized density were used for plotting the data (Fig. 2a-c). It is evident from these maps that the high degree of small scale variation would not facilitate the geographical definition of strata boundaries using the defined limits of normalized density. The plots do reveal, however, some persistent features:

1. Low relative densities in the Scotian Gulf.
2. High relative densities along the northern edge of Western Bank.
3. High relative densities in the northeast portion.
4. Low relative densities on southern Banquereau Bank.
5. Variable relative densities on Western Bank.

Based on these observations, three strata were defined, subjectively in each of Sub-Division 4Vs and Division 4W, respecting the Division boundaries (Fig. 2d).

Allocation

The method used to define strata limits has an associated allocation scheme, but since strata boundaries were defined subjectively, an alternative approach was used. Hansen et al. (1953, p. 215) note that an approximately optimal allocation scheme, when means and standard deviations within strata are proportionately related, is to allocate in proportion to the population size in each stratum. The relationship between the estimated means and standard deviations from the six strata over all years appears linear with an intercept near the origin (Fig. 3). The proportion of the total population in each stratum varies from year to year and there is a tendency for larger values in strata 1, 4 and 6 (Table 3). The average pattern was used for allocation (Table 4).

Evaluation

The method described by Sukhatme and Sukhatme (1970, p. 91) was used to estimate the gain in precision due to stratification, relative to simple random sampling. For arbitrary allocation of stations among strata the difference between the variance from simple random sampling and the variance from stratified sampling is given by:

$$(1)\quad V(\overline{Y}_R) - V(\overline{Y}_S) \cong \sum_h \left(\frac{1}{n} - \frac{W_h}{n_h}\right) W_h S_h^2 + \left(\frac{1}{n} - \frac{1}{N}\right) \sum_h W_h (\overline{Y}_h - \overline{Y})^2$$

where N_h = Number of sampling units in stratum h

$N = \sum_h N_h$

$W_h = N_h/N$

$\overline{Y}_h$ = Population mean in stratum h

$\overline{Y} = \sum_h W_h \overline{Y}_h$

S_h^2 = Variance in stratum h

$V(\overline{Y}_R)$ = Variance for population mean from simple random sampling

$V(\overline{Y}_S)$ = Variance for population mean from stratified sampling

n_h = Sample size in stratum h

$n = \sum_h n_h$

The first term on the right, which we will call the allocation component, can be negative, positive or zero depending on the allocation scheme used. The second term on the right, which we refer to as the strata component, is always greater than or equal to zero.

The quantity in Equation 1 can be estimated by:

$$(2)\quad \text{Est.}[V(\overline{Y}_R)-V(\overline{Y}_S)] = \sum_h \left(\frac{1}{n} - \frac{W_h}{n_h}\right) W_h s_h^2 + \frac{N-n}{(N-1)n} \sum_h W_h (\overline{y}_h - \overline{y}_{st})^2 - \sum_h W_h (1-W_h) s_h^2/n_h$$

It is evident that Equation 2 retains the decomposition to allocation and strata components. It is more informative to examine the relative gain in efficiency given by:

(3) Efficiency = [Est. $[V(\bar{Y}_R) - V(\bar{Y}_S)]$ / Est. $V(\bar{Y}_R)$] x 100%

Further, due to the additive nature of the allocation and strata components, the percent efficiency gain due to each can be extracted.

For each of the domains of study, 4Vs, 4W and 4VsW, Equations 2 and 3 were applied to both the present stratification scheme and the alternative sampling design. Two cases are presented for the alternative sampling design a) with the allocation based on the existing surveys and b) using the same strata means and variances, but changing the allocation to the one prescribed by the proportions in Table 4. It should be noted that the domains for the present and new sampling designs are not exactly equivalent since the new strata boundaries follow Division lines. The current strata borders overlap the Division boundaries. This difference should not greatly affect comparisons.

The most prominent feature in Figs. 4, 5 and 6 is the large penalty incurred by the present sampling scheme on the allocation component. This effect is more pronounced in Sub-Division 4Vs. In general, the gains due to the strata component are moderate in comparison to the magnitude of the allocation effect. The strata component is defined to be a positive quantity but negative estimates can occur. These occurrences appear to be due to very imprecise estimates of s_h^2 based on small sample sizes. The new stratification scheme resulted in very small gains in efficiency relative to simple random sampling. Application of the new allocation scheme resulted in measurable gains in some years, but was also susceptible to greater penalties in other years.

Discussion

A number of studies have examined the level of precision for abundance indices from bottom trawling surveys (Grosslein, 1971; Jones and Pope, 1972; Pennington and Grosslein, 1978). In this study we compare the performance of a depth based stratified random design with one based on historical distribution patterns using simple random sampling as a reference. Recognizing that the stratification and allocation schemes derived from the historical data are applied to the same data, it may be expected that the variances for these designs are somewhat underestimated. The performance is evaluated with respect to a single objective, the precision of abundance estimates.

The most striking feature of the comparisons is that the present stratified random design is considerably worse than simple random sampling for the large majority of the cases. It has been noted that, due to the allocation component, arbitrary allocation schemes may result in large losses of efficiency (Sukhatme and Sukhatme, 1970). Primarily due to the large number of strata which restricts the options available for allocation, this phenomenon seems to have occurred here. Even the use of historical data to assign the allocation of stations resulted in some years with substantial penalties in efficiency. These years appear to be occasions when the proportion of the total estimate caught in each strata (Table 3) substantially deviated from the average proportion used to assign the sets.

When the number of strata are increased then the gain in efficiency due to stratification is expected to be greater. It is notable in these comparisons that a reduction from 24 strata to 6 strata did not result in large decreases in this component. This is due primarily to the very small gains accounted for by the strata component. This observation supports the statement by Cochran (1977) regarding expected gains by increasing the number of strata.

The poor performance of the present sampling design and the mediocre performance of the alternative sampling design are an indication of the highly variable nature of trawl catches which has been noted previously. It has also been noted that geographical stratification rarely produces significant gains in efficiency. Based on the results, it could be concluded that, with respect to the precision of abundance estimates, it would be difficult to do much better than simple random sampling for the summer surveys on the eastern Scotian Shelf. If it was considered useful to stratify in order to spread out stations, then the number of strata should be reduced to allow sampling proportional to the size of the strata. The allocation component for proportional sampling is zero, ensuring that efficiency is not worse than simple random sampling. From the results presented here, its appears that a reduction in the number of strata would not result in significant decreases in the between-strata component.

References

Cochran, W.G. 1977. Sampling Techniques. John Wiley and Sons Inc., New York. 428 p.

Doubleday, W.G. [ed.] 1981. Manual on groundfish surveys in the Northwest Atlantic. N.A.F.O. Sci. Coun. Studies, 2: 55 p.

Gagné, J.A., A.F. Sinclair, and L. Currie. MS 1983. Status of the 4VsW Cod Stock Complex. CAFSAC Res. Doc. 83/56.

Gagné, J.A., A.F. Sinclair, and C. Dale. MS 1984. The 1984 Assessment of 4VsW Cod -- A Completely Revised Procedure. CAFSAC Res. Doc. 84/78.

Grosslein, M. D. 1969a. Groundfish survey program of BCF Woods Hole. Comm. Fish. Rev., 31: 22-30.

Grosslein, M. S. 1969b. Groundfish Survey Methods. U.S. Bur. Comm.

Grosslein, M.D. MS 1971. Some Observations on Accuracy of Abundance Indices Derived from Research Vessel Surveys. ICNAF Redbook 1971, Part III: 249-266.

Halliday, R.G. and P.A. Koeller. 1981. A History of Canadian Groundfish Trawling Surveys and Data Useage in ICNAF Divisions 4TVWX. p. 27-41. In: W.G. Doubleday and D. Rivard [eds.], Bottom Trawl Surveys. Can. Spec. Publ. Fish. Aquat. Sci. 58: 273 p.

Hansen, M.H., W.N. Hurwitz, and W.G. Madow. 1953. Sample Survey Methods and Theory. John Wiley and Sons. Vol. 1.

ICNAF. 1970. Report of Standing Committee on Research and Statistics. App V. Report of Ad Hoc Working Group on Coordinated Groundfish Surveys. ICNAF Redbook 1975: 81-83.

Jones, B.W. and J.G. Pope. MS 1972. A Groundfish Survey of Faroe Bank. ICNAF Res. Doc. No. 64, Serial No. 2783.

Pennington, M.R. and M.D. Grosslein. MS 1978. Accuracy of Abundance Indices Based on Stratified-Random Surveys. ICNAF Res. Doc. No. 77, Serial No. 5264.

Scott, J.S. and S. Gavaris. MS 1985. (This Meeting)

Sinclair, A. and S. Gavaris. MS 1985. Sequential Population Analysis of 4VsW Cod Following the 1984 Fishery. CAFSAC Res. Doc. 85/48.

Sukhatme, R.V. and B.V. Sukhatme. 1970. Sampling Theory of Surveys with Applications. Food and Agriculture Organization, Rome, Second Edition.

Table 1. Sampling schedule for strata on the Scotian Shelf and Bay of Fundy. The number of random stations is indicated in the column labelled "n".

Scotian Shelf				Bay of Fundy			
(Sub-) Division	Stratum No.	Area (n.m.)	n	(Sub-) Division	Stratum No.	Area (n.m.)	n
	40	924	3		70	920	2
4Vn	41	1,000	3		71	1,004	2
	42	1,437	3		72	1,249	2
					73	265	2
	43	1,318	4		74	161	2
	44	3,925	4		75	156	2
	45	1,023	4	4X	76	1,478	2
4Vs	46	491	3		77	1,232	2
	47	1,616	4		78	233	3
	48	1,449	4		80	655	4
	49	144	2		81	1,875	4
	50	383	3		82	1,042	2
	51	147	2		83	532	2
	52	345	2		84	2,264	3
					85	1,582	3
	53	259	3		90	601	3
	54	499	3		91	687	3
	55	2,122	7		92	1,086	3
	56	955	6		93	533	3
	57	811	2		94	417	2
	58	658	3		95	584	2
4W	59	3,148	4				
	60	1,344	2				
	61	1,154	2				
	62	2,116	4				
	63	302	2				
	64	1,297	5				
	65	2,383	5				
	66	226	3				

Table 2. Cumulative square root of the frequency, showing the upper and lower limits of the density (normalized) for each of the strata. Strata A and B can not be differentiated without using a finer resolution for the classes. Classes 9-20 are aggregated in this table for convenience.

	Density (normalized) Limits		Cumulative Square Root Frequency			
Class	Lower	Upper	1970-74	1975-79	1980-84	Strata
1	0.00	0.75	17.4	17.0	17.9	A + B
2	0.75	1.50	23.4	23.8	24.5	C
3	1.50	2.25	27.4	29.0	28.3	D
4	2.25	3.00	31.0	31.7	31.1	
5	3.00	3.75	33.3	34.7	33.3	E
6	3.75	4.50	35.0	36.7	35.6	
7	4.50	5.25	36.0	38.9	37.8	
8	5.25	6.00	37.4	39.9	38.8	
9-20	6.00	15.00	43.8	48.1	48.7	F

Table 3. Proportion of the total population in each strata (new boundaries).

	STRATA					
Year	1	2	3	4	5	6
1970	0.559	0.043	0.017	0.022	0.178	0.181
1971	0.630	0.008	0.021	0.076	0.203	0.062
1972	0.635	0.020	0.042	0.160	0.083	0.060
1973	0.131	0.023	0.020	0.094	0.004	0.728
1974	0.357	0.027	0.113	0.104	0.077	0.321
1975	0.263	0.122	0.139	0.098	0.105	0.274
1976	0.195	0.002	0.033	0.341	0.160	0.269
1977	0.315	0.096	0.061	0.138	0.074	0.315
1978	0.046	0.034	0.033	0.119	0.640	0.127
1979	0.048	0.107	0.045	0.333	0.316	0.151
1980	0.206	0.143	0.005	0.332	0.195	0.118
1981	0.384	0.083	0.016	0.274	0.146	0.096
1982	0.365	0.057	0.069	0.436	0.027	0.045
1983	0.298	0.081	0.150	0.380	0.051	0.040
1984	0.675	0.025	0.022	0.066	0.179	0.033

Table 4. Sampling schedule for the new stratification scheme in Sub-Division 4Vs and Division 4W. The proportion of random stations is indicated in the column labelled "p".

(Sub-) Division	Stratum No.	Area (n.m.)	p
4Vs	1	3,600	0.328
	2	2,280	0.057
	3	3,690	0.053
4W	4	7,465	0.201
	5	2,600	0.146
	6	8,400	0.214

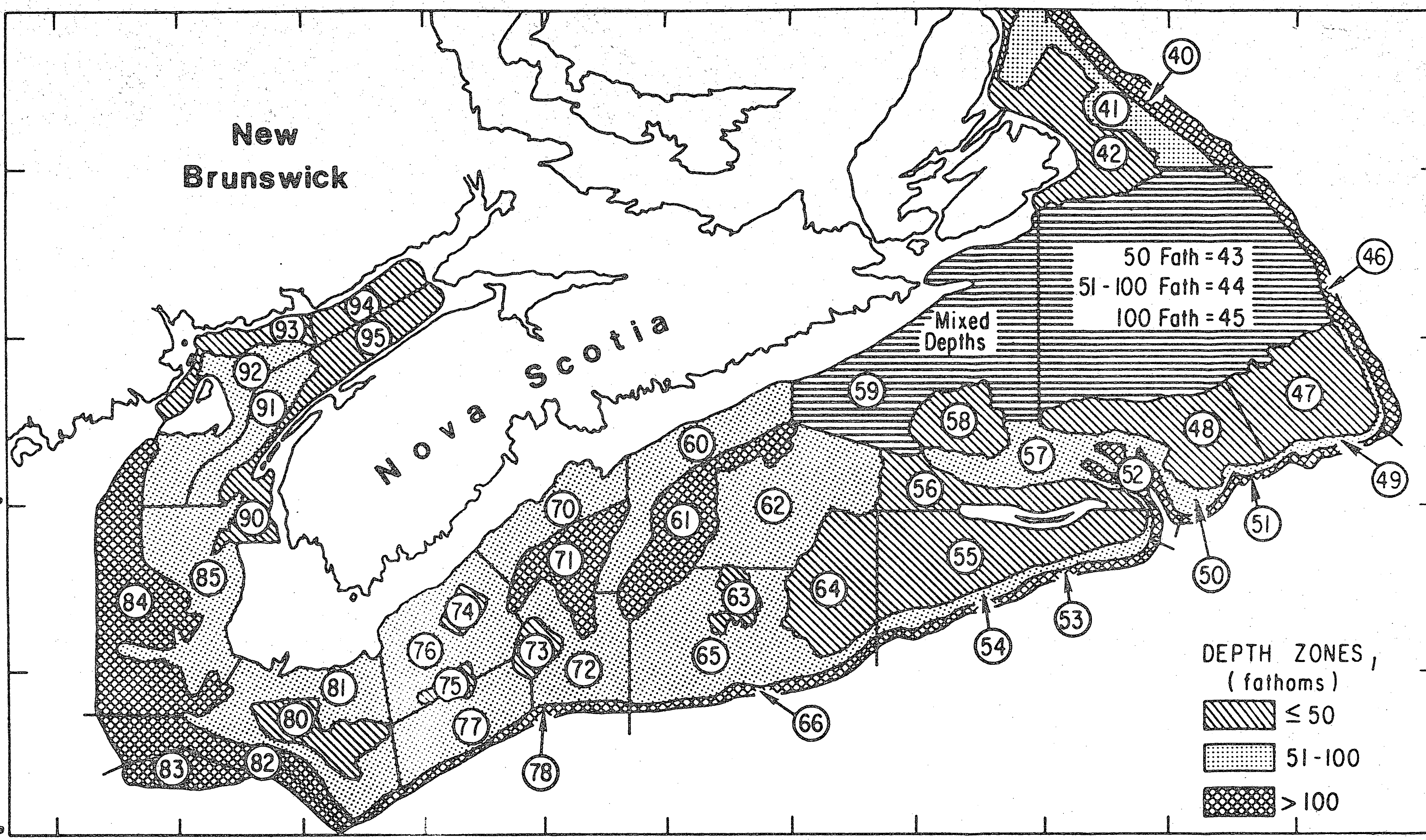

Fig. 1. Stratification and numbering system for Divisions 4V, 4W and 4X adopted in 1970.

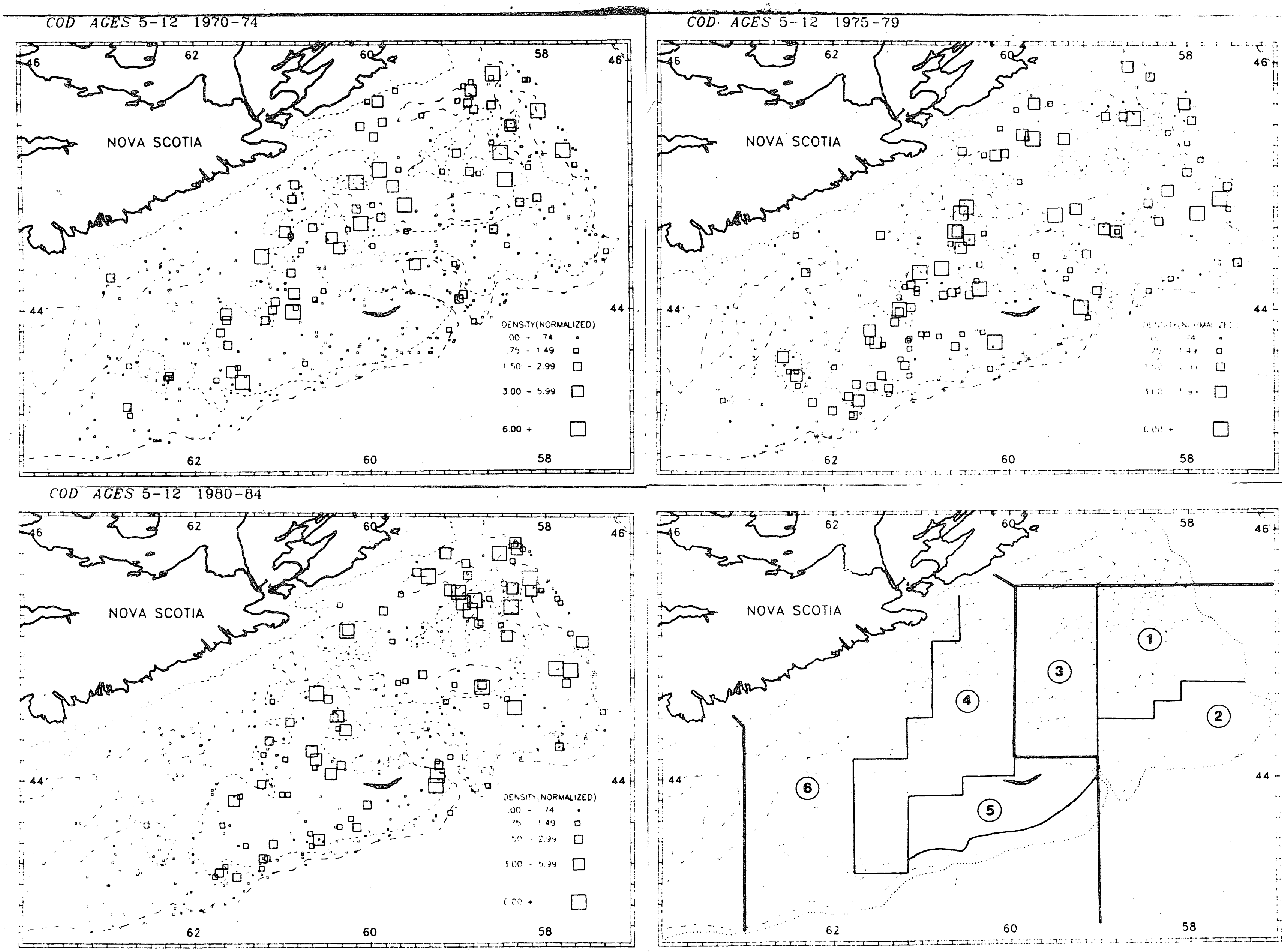

Fig. 2. Normalized density plots for the three 5-year periods and the alternate stratification design. (----50 fathoms; --- 100 fathoms; ...... 200 fathoms)

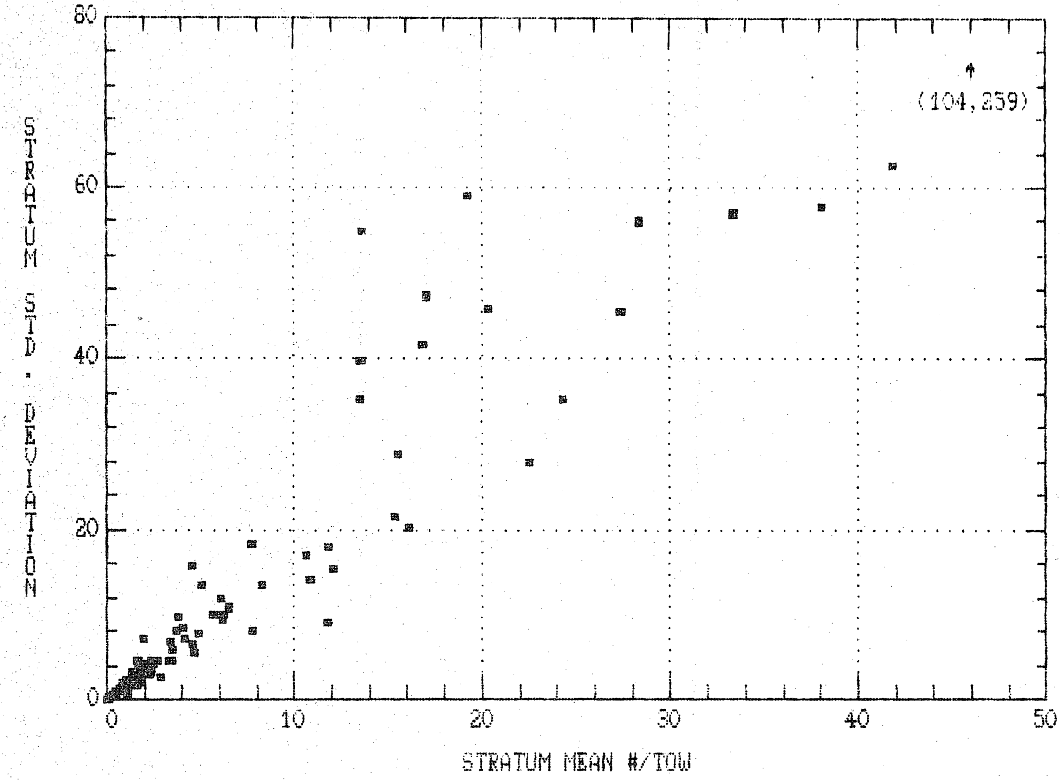

Figure 3: Comparison of Stratum means and standard deviations. 4VsW Cod, 1970-1984. New strata boundaries (see text).

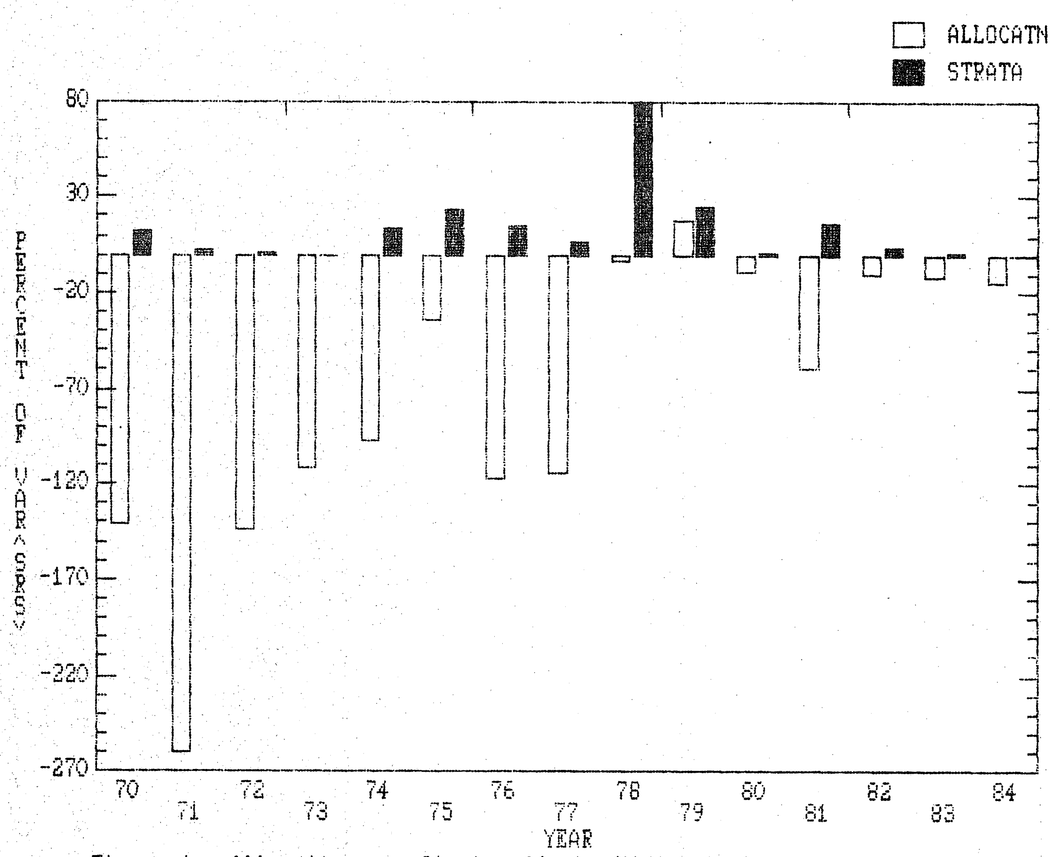

Figure 4a: Allocation vs. Strata effect. 4VsW cod. Current strata boundaries.

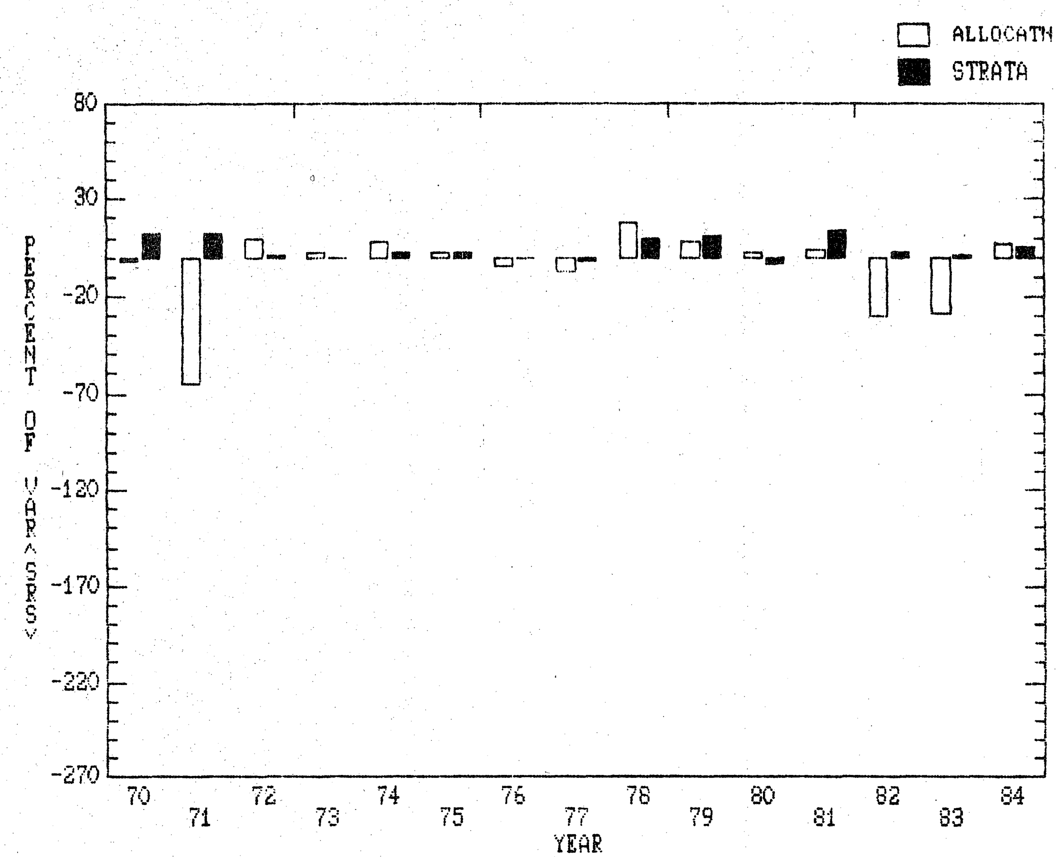

Figure 4b: Allocation vs. Strata effect. 4VsW Cod. New strata boundaries.

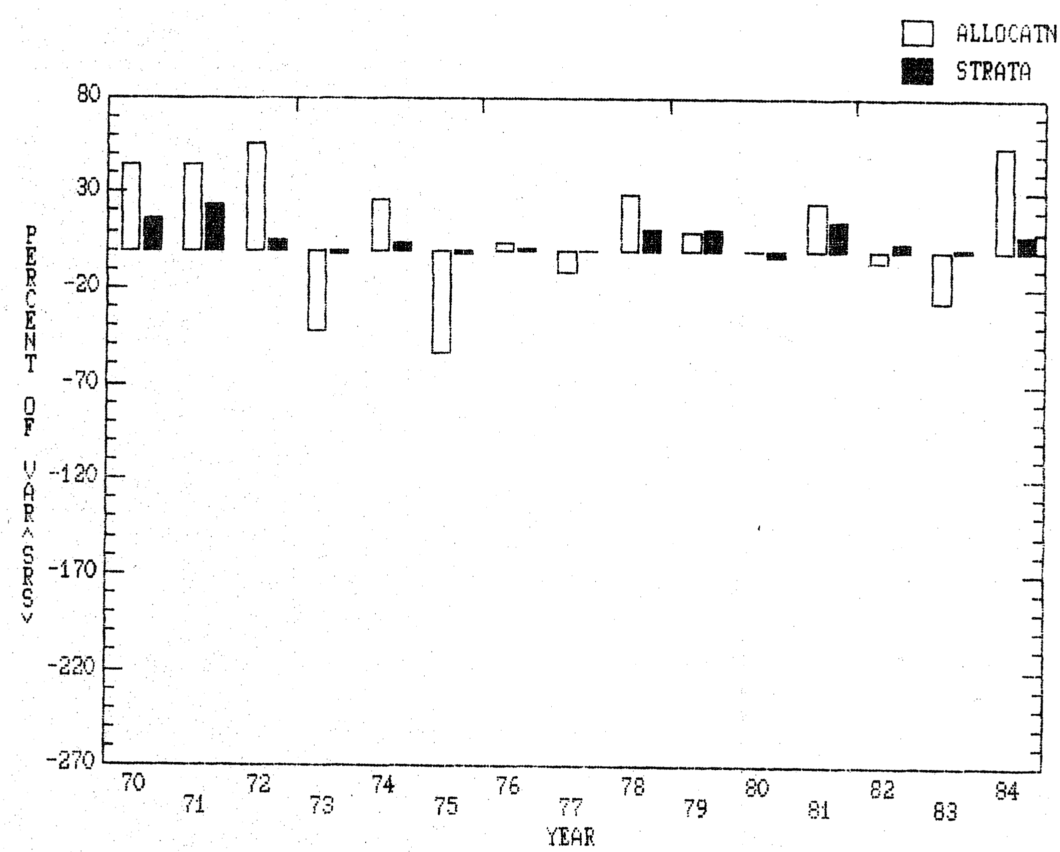

Figure 4c: Allocation vs. Strata effect. 4VsW Cod. New strata-allocation by stratum total.

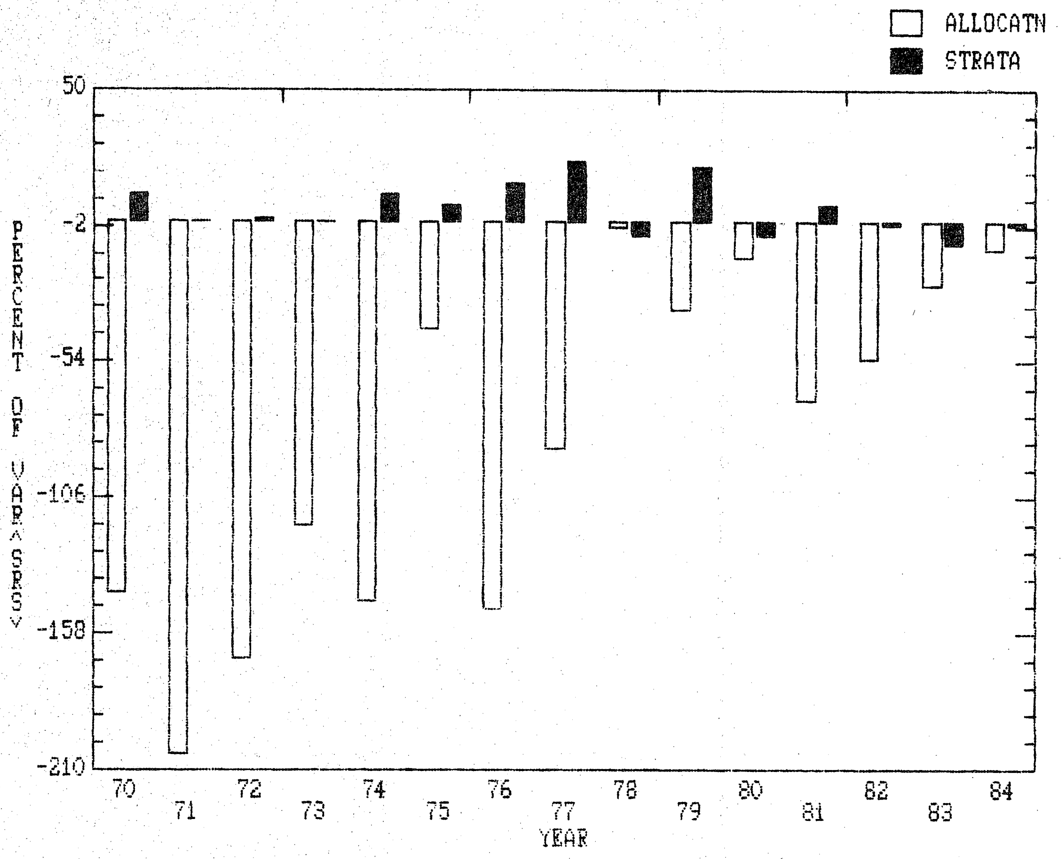

Figure 5a: Allocation vs. Strata effect. 4Vs Cod, Current strata boundaries.

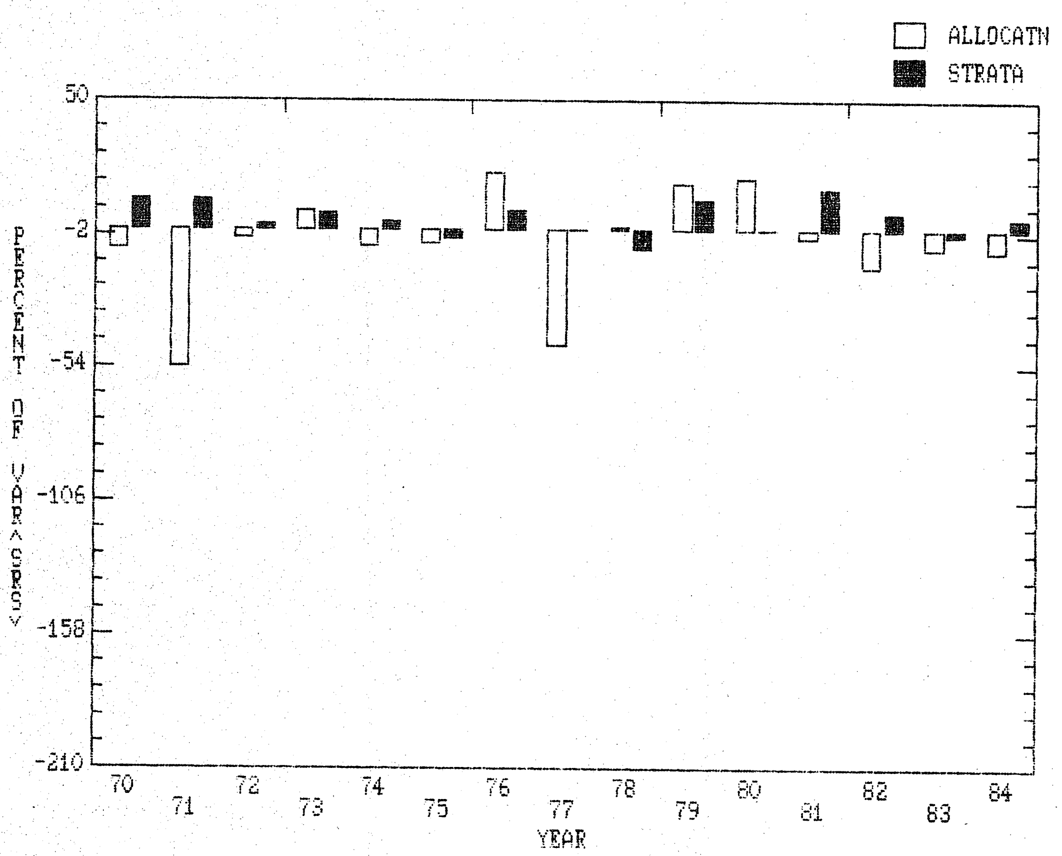

Figure 5b: Allocation vs. Strata effect. 4Vs Cod, New strata boundaries.

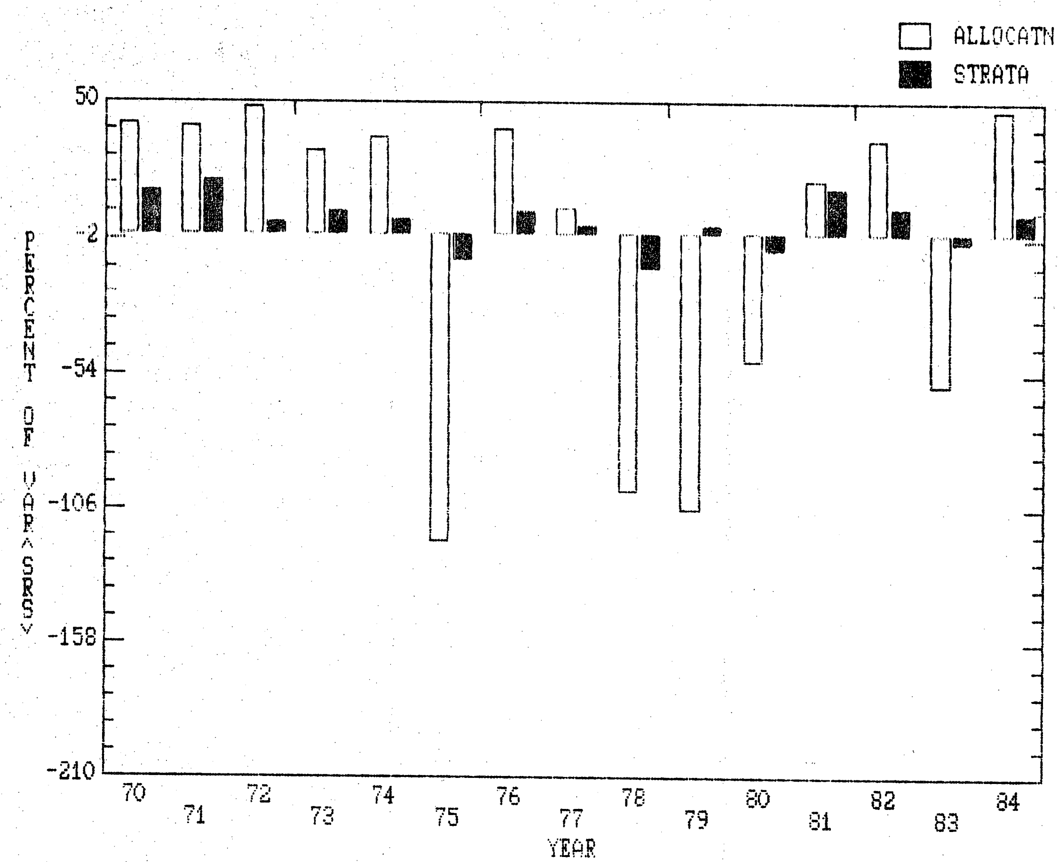

Figure 5c: Allocation vs. Strata effect. 4Vs Cod. New strata-allocation by stratum total.

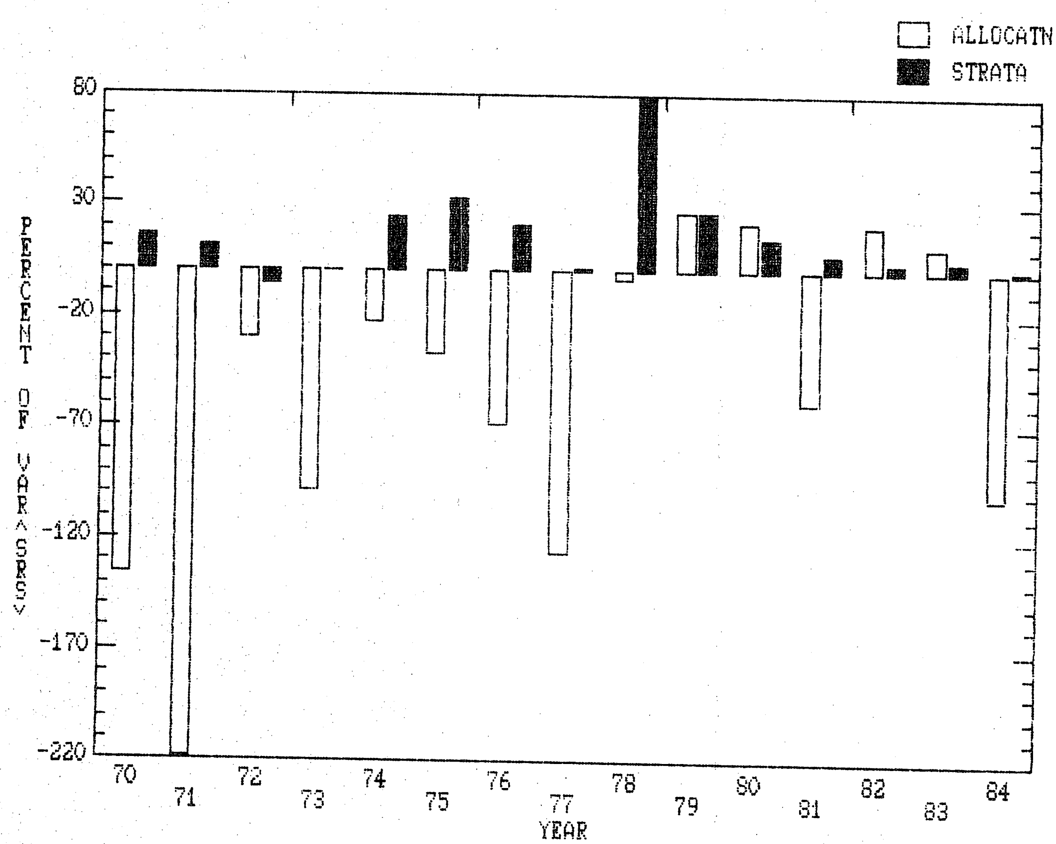

Figure 6a: Allocation vs. Strata effect. 4W cod. Current strata boundaries.

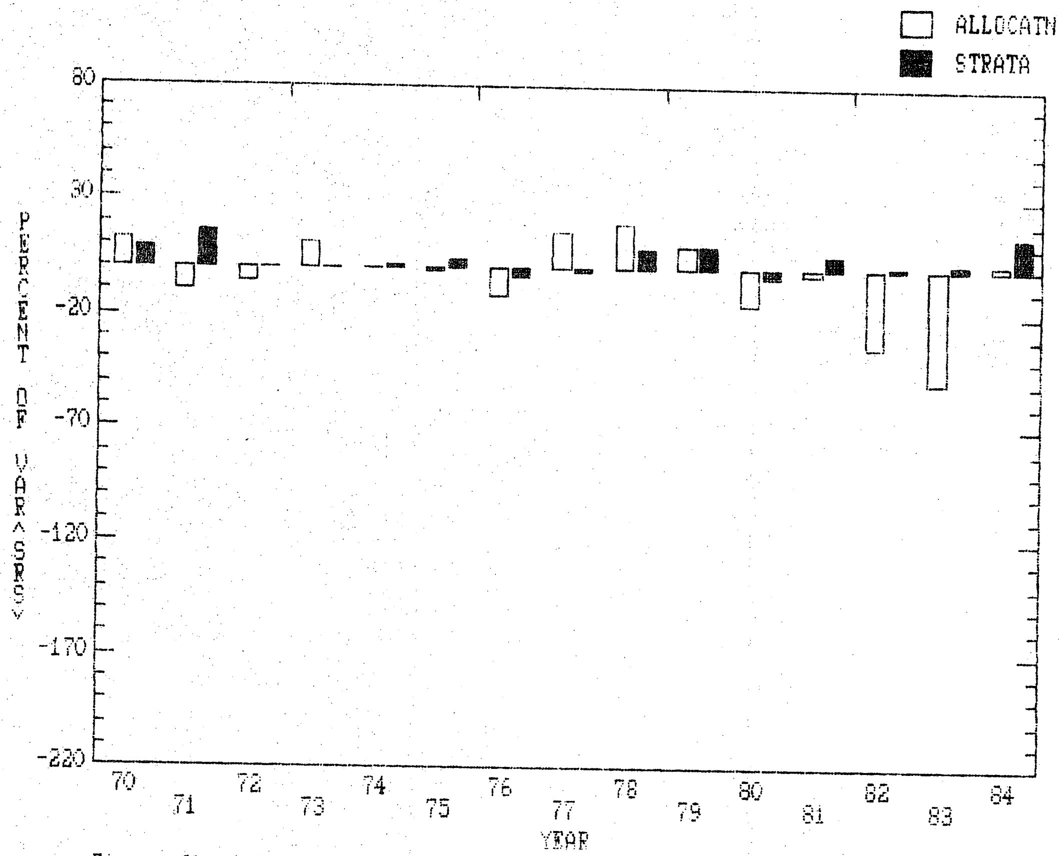

Figure 6b: Allocation vs. Strata effect. 4W Cod. New Strata boundaries.

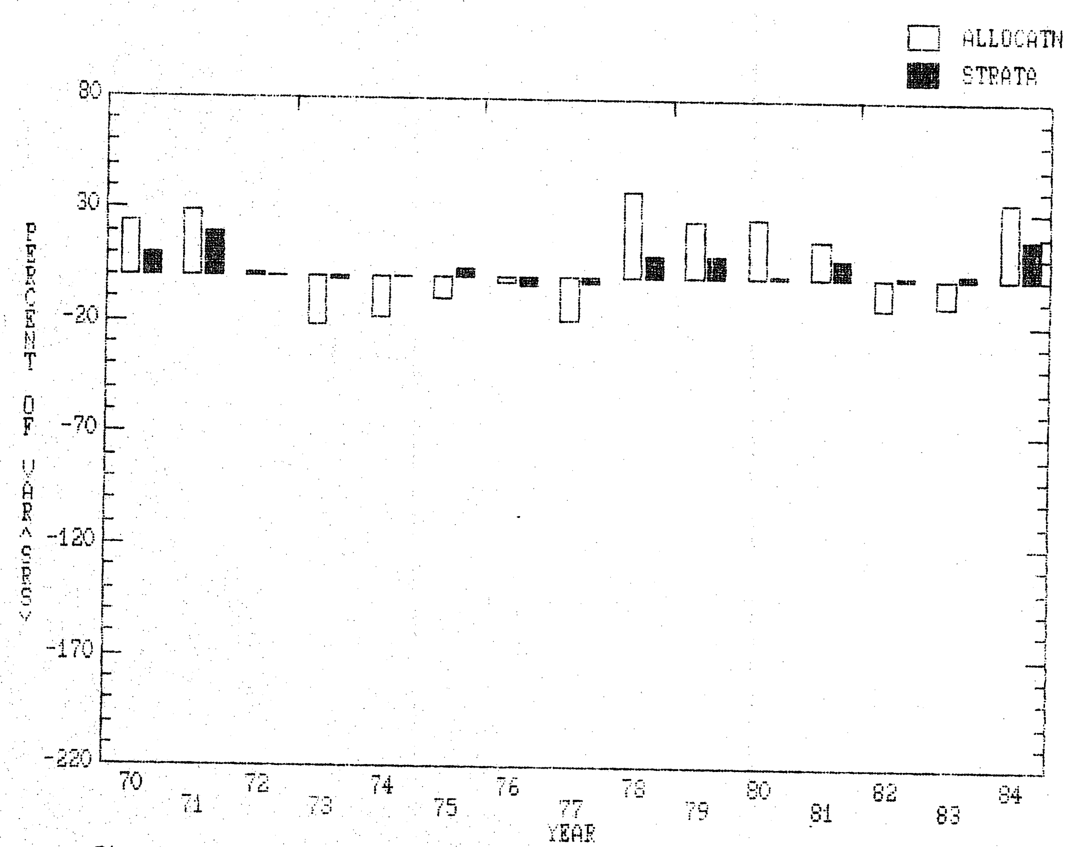

Figure 6c: Allocation vs. Strata effect. 4W Cod. New strata-allocation by stratum total