This Paper Not to be Cited Without Prior Reference to the Author

International Commission for
the Exploration of the Sea

C.M. 1977/F:32
Demersal Fish
(Northern) Committee

An Analysis of Factors Influencing Catch Rates of Juvenile Roundfish from North Sea Young Herring Surveys

W.H. Marshall
Marine Fish Division
Bedford Institute of Oceanography
Dartmouth, Nova Scotia
Canada

Introduction

At the meeting of the Gadoid I-Group Working Group which was held in Ymuiden from 24-26 May 1977, there was some discussion of the problems associated with the analysis of juvenile survey data. The major problems arise from the large between-ship variability in reported catch rates. A traditional approach to solving this problem has been to try to reduce the variability by standardizing gear. At best this is only a partial solution since differences in catch rate are still bound to occur because of the differences in fishing power of the boats involved in the survey. Unless vessel effects are somehow removed, the calculated abundance estimates will be biased and the confidence intervals will be wider than they need to be.

The analysis of variance (ANOVA) is a statistical technique for partitioning the total variation of a dependent variable (catch rate) into a number of component parts. These parts can be identified with particular factors (vessel, year, depth, etc.) that are thought to influence the dependent variable. Any of the variation which is not accounted for by these factors is called the residual and is an estimate of the variation inherent in the sampling method (trawling). Statistical tests are available for determining which of the factors contribute significantly to the overall variation.

This paper presents the results of an analysis of variance which was performed on the data collected for juvenile roundfish in the North Sea.

The Data

The dependent variables used in the analysis consisted of the hourly catch rates for I and II group cod, haddock, and whiting. The rates were reported by individual haul but for some hauls there was no breakdown of the total catch rate into age groups. These hauls were treated as missing values and are not included in the analysis.

The independent variables (factors, main effects) investigated in the analysis were year, vessel, date, time of day, location, depth, bottom temperature and bottom salinity. Since temperature and salinity data were only available for the years 1970 through 1974, these were the only years included in the analysis. Hauls for which any of the other independent variables were not measured were excluded from the analysis.

The Analysis

A basic assumption of the analysis of variance is that the dependent variables are independently and normally distributed with a common variance. Histograms (Figures 1 and 2) and distributional statistics (Table 1) of the catch rates both indicate that the data is not normally distributed. In fact, as was noted at the working group meeting, it seems to follow a negative binomial distribution. To bring the data more in line with the theoretical requirements, all the catch rates were transformed by the function $f(x) = \ln (x + 1)$. The resulting variables are still not normally distributed, due mainly to the large number of hauls in which zero fish were caught, but they are closer to normality than the original variables. This can be seen by comparing the histograms and distributional statistics of the original and transformed variates. By examining the catch rates for individual vessels it was also seen that the transformed variables had a more constant variance than the original variables. For these reasons the transformed variables were used in the remainder of the analysis.

When the data is balanced (same number of observations for each combination of factors) the analysis is relatively simple but when it is unbalanced, as it is in this case, it is not feasible to perform the analysis without a computer. The ANOVA's presented in this paper were generated using the SPSS statistical package on an IBM computer. To some extent the limitations of the program governed the order in which the various factors were investigated and the way they were intorduced into the model (i.e. as factors or covariates). This was expecially true for the analysis of location. Ideally, the individual square numbers should have been introduced into the model as main effects. Unfortunately, this resulted in a design matrix which was much too large for the program to handle and it therefore became necessary to group the hauls into seven rather arbitrary areas (Figure 3), which were used instead of square numbers in the analysis. Similarily depth, temperature and salinity should have been main effects. However, SPSS requires considerably more space to analyse the main effect than a covariate. Therefore, these environmental factors were put into the model as covariates and were examined only for linear and quadratic significance. The final ANOVA tables for each of the three species and two age groups are presented as Tables 2 through 7.

The main problems encountered in this analysis were of a strictly computational nature. It is impossible to perform the analysis without the aid of a computer and yet the computer system forces comprimises to be made in the method of analysis. Even though significant results were obtained from this analysis it would be worth while to look for a program which would allow a more accurate analysis to be made.

The last step in the analysis consisted of doing a multiple classification analysis to determine the average catch rate for each year after compensating for the fact that several different vessels were participating in the survey (Figures 4 through 9). The analysis also gave the average catch rate of each vessel after correcting for the different years (Table 8).

Results

The factors salinity, date, and time were all tested and found to be non-significant. Salinity and date were put into the model as covariates. When a factor is used as a covariate, the analysis performed is similar to that done in a regression analysis and only linear relationships between the dependent variable and the covariates will be detected. Therefore, it is still possible that salinity and date do have a significant effect on catch rates if that effect is non-linear. Time was analysed to see if there was a significant diurnal variation in catch rates. Hauls were divided into two groups depending whether they occurred during the day (between 07:10 and 16:55) or at night. The resulting grouping was treated as a main effect but was not found to be significant.

The variables year, vessel, and area were all analysed as main effects and found to be significant. It is not surprising that any of these effects are significant and in fact, it would be more surprising if they were not. What is more interesting is that many of the interactions between main effects are also significant. The year-vessel interaction is significant in all cases and indicates that the relative fishing power between vessels is changing from one year to the next. A possible cause of this would be vessels changing gear type or personnel from one year to the next. Similarly the significant year-area interactions indicate that the relative concentration of fish between areas is changing. A significant vessel-area interaction indicates that the relative catch rates between vessels are changing as they move from one area to the next. This may be the case if, for instance, one area had predominantly deep water with a smooth bottom giving the larger vessels an advantage while a second area had shallow water with an irregular bottom giving more of an advantage to the smaller boats.

The depth, temperature, and depth-temperature interactions were analysed as covariates and found to be highly significant in all cases except for II-Cod. A possible explanation for why it is only the catch rates for II-Cod which are not affected by depth and temperature, is that at age 2, cod are starting to go to the bottom and at the time of the survey they would be fairly evenly distributed over depths since the entire population of II-Cod would not be on the bottom yet. The calculated slopes for the covariates are shown on the bottom of the ANOVA tables and are seen to be relatively consistent from one species to the next.

A multiple classification analysis table was printed as part of the output from the ANOVA. Figures 4 through 9 show the average catch rates for 1970 through 1974 for each species and age group. These figures show that in some cases correcting the catch rates by eliminating the biasing effect of vessel and area fished makes a significant difference to the abundance estimates which are generated. Similarly, the average catch rates for each vessel are different depending on whether or not they are corrected for the effects of year and area. Unfortunately, the corrected catch rates for vessels represent only the average differences which have existed in the years 1970 through 1974. They cannot be used as correction factors for future catches because of the significant interactions between vessel and years. If it is possible to remove the vessel-year interaction from future surveys, perhaps by gear standardization, it would be possible to calculate correction factors which would permit the tows from any given year to be expressed in terms of a standard vessel without necessarily having to repeat this analysis. This would make it possible to compare hauls from different vessels without having to worry about whether differences in catch rates are caused by actual differences in abundance or differences in fishing power of the vessels. It therefore seems reasonable that one of the objectives of future surveys should be to try to keep the relative catch rates from one vessel to another as constant as possible.

Table 1. Distributional statistics of research catch rates.

	Number of Hauls	Before Transformation				After Transformation*			
		Mean	Standard Deviation	Skewness	Kurtosis	Mean	Standard Deviation	Skewness	Kurtosis
I - Cod	755	43.9	155.9	7.7	73.5	1.96	1.73	.77	- .06
II - Cod	755	23.1	88.0	10.3	136.5	1.69	1.49	.87	.35
I - Haddock	625	688.4	1523.0	4.4	26.5	4.06	2.74	-.03	-1.29
II - Haddock	625	263.4	887.2	8.1	79.3	3.27	2.35	.15	-1.00
I - Whiting	771	423.2	1078.3	6.0	47.2	4.26	2.10	-.10	- .62
II - Whiting	771	313.3	1697.8	12.9	194.9	3.17	2.21	.39	- .46

* the catch rates were transformed by the function F(x) = ln(x+1).

Table 2. Analysis of variance table and covariate regression parameters for I - Cod.

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	733.2	25	29.3	17.468	0.001
Year	236.3	4	59.0	35.196	0.001
Vessel	278.5	15	18.5	11.058	0.001
Area	90.7	6	15.1	9.010	0.001
Covariates	38.1	3	12.7	7.577	0.001
Depth	21.0	1	21.0	12.559	0.001
Temperature	13.5	1	13.5	8.059	0.005
Depth x Temperature	16.3	1	16.3	9.729	0.002
2-Way Interactions	289.4	71	4.0	2.428	0.001
Year x Vessel	48.8	12	4.0	2.424	0.005
Year x Area	84.7	21	4.0	2.404	0.001
Vessel x Area	93.6	38	2.4	1.468	0.037
Explained	1060.8	99	10.7	6.832	0.001
Residual	1047.7	624	1.6		
Total	2108.6	723	2.9		

Covariate	Beta
Depth	-0.033
Temperature	-0.306
Depth x Temperature	0.005

Table 3. Analysis of variance table and covariate regression parameters for II - Cod.

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	283.2	25	11.3	7.172	0.001
Year	37.8	4	9.4	5.990	0.001
Vessel	105.6	15	7.0	4.459	0.001
Area	84.0	6	14.0	8.869	0.001
Covariates	3.8	3	1.2	0.820	0.999
Depth	0.0	1	0.0	0.008	0.999
Temperature	0.1	1	0.1	0.084	0.999
Depth x Temperature	0.0	1	0.0	0.019	0.999
2-Way Interactions	211.9	71	2.9	1.890	0.001
Year x Vessel	51.0	12	4.2	2.692	0.002
Year x Area	46.4	21	2.2	1.400	0.110
Vessel x Area	97.9	38	2.5	1.632	0.011
Explained	499.1	99	5.0	3.191	0.001
Residual	985.7	624	1.5		
Total	1484.9	723	2.0		
Covariates	Beta				
Depth	0.001				
Temperature	-0.030				
Depth x Temperature	0.000				

Table 4. Analysis of variance table and covariate regression parameters for I - Haddock.

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	2489.1	25	99.5	31.533	0.001
Year	438.8	4	109.7	34.751	0.001
Vessel	118.0	15	7.8	2.493	0.002
Area	1150.7	6	191.7	60.744	0.001
Covariates	79.6	3	26.5	8.410	0.001
Temperature	70.9	1	70.9	22.477	0.001
Depth	42.5	1	42.5	13.491	0.001
Depth x Temperature	40.2	1	40.2	12.732	0.001
2-Way Interactions	418.2	68	6.1	1.948	0.001
Year x Vessel	81.2	13	6.2	1.979	0.021
Year x Area	103.5	19	5.4	1.726	0.029
Vessel x Area	142.7	36	3.9	1.256	0.150
Explained	2987.0	96	31.1	9.855	0.001
Residual	1622.9	514	3.1		
Total	4610.0	610	7.5		

Covariate	Beta
Temperature	1.299
Depth	0.059
Depth x Temperature	-0.009

Table 5. Analysis of variance table and covariate regresssion parameters for II - Haddock.

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	1880.0	25	75.2	36.121	0.001
Year	522.4	4	130.6	62.742	0.001
Vessel	158.2	15	10.5	5.067	0.001
Area	754.5	6	125.7	60.404	0.001
Covariates	55.9	3	18.6	8.965	0.001
Temperature	47.1	1	47.1	22.631	0.001
Depth	22.3	1	22.3	10.729	0.001
Depth x Temperature	20.7	1	20.7	9.966	0.002
2-Way Interactions	379.1	68	5.5	2.678	0.001
Year x Vessel	109.6	13	8.4	4.053	0.001
Year x Area	70.6	19	3.7	1.786	0.022
Vessel x Area	143.0	36	3.9	1.909	0.001
Explained	2315.1	96	24.1	11.584	0.001
Residual	1070.0	514	2.0		
Total	3385.2	610	5.5		

Covariate	Beta
Temperature	1.059
Depth	0.042
Depth x Temperature	-0.007

Table 6. Analysis of variance table and covariate regression parameters for I - Whiting.

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	881.1	26	33.8	11.817	0.001
Year	112.4	4	28.1	9.806	0.001
Vessel	297.6	16	18.6	6.487	0.001
Area	299.3	6	49.8	17.395	0.001
Covariates	83.0	3	27.6	9.655	0.001
Temperature	44.9	1	44.9	15.664	0.001
Depth	31.1	1	31.1	10.864	0.001
Depth x Temperature	42.4	1	42.4	14.788	0.001
2-Way Interactions	498.9	72	6.9	2.416	0.001
Year x Vessel	146.6	10	14.6	5.115	0.001
Year x Area	77.8	21	3.7	1.293	0.171
Vessel x Area	186.3	41	4.5	1.585	0.013
Explained	1463.1	101	14.4	5.051	0.001
Residual	1898.6	662	2.8		
Total	3361.8	763	4.4		

Covariate	Beta
Temperature	0.581
Depth	0.043
Depth x Temperature	-0.008

Table 7. Analysis of variance table and covariate regression parameters for II - Whiting.

Source of Variation	Sum of Squares	DF	Mean Square	F	Significance of F
Main Effects	1161.6	26	44.6	15.660	0.001
Year	395.2	4	98.8	43.635	0.001
Vessel	154.5	16	9.6	3.386	0.001
Area	310.7	6	51.7	18.153	0.001
Covariates	106.9	3	35.6	12.490	0.001
Temperature	97.3	1	97.3	34.114	0.001
Depth	22.6	1	22.6	7.954	0.005
Depth x Temperature	22.8	1	22.8	9.024	0.005
2-Way Interactions	565.2	72	7.8	2.752	0.001
Year x Vessel	84.8	10	8.4	2.976	0.001
Year x Area	180.2	21	8.5	3.009	0.001
Vessel x Area	160.0	41	3.9	1.368	0.066
Explained	1833.8	101	18.1	6.364	0.001
Residual	1888.7	662	2.8		
Total	3722.5	763	4.8		

Covariate	Beta
Temperature	0.855
Depth	0.037
Depth x Temperature	-0.006

Table 8. Average corrected catch rates for each vessel.

Vessel	I - Cod	II - Cod	I - Haddock	II - Haddock	I - Whiting	II - Whiting
Ernest Holt (England)	4.00	3.95	35.23	15.61	9.70	7.94
Anton Dohrn (FRG)	1.36	2.42	55.83	18.49	33.12	16.46
Tridens (The Netherlands)	5.23	5.82	76.48	28.37	90.84	34.52
Willem Beukelsz (The Netherlands)	4.87	6.39	-	-	58.74	9.28
Scotia (Scotland)	4.70	5.17	28.08	11.30	34.16	1.48
Dana (Denmark)	28.67	6.77	52.52	63.72	272.14	36.34
G.O. Sars (Norway)	2.90	2.49	74.94	68.41	171.43	94.58
Explorer (Scotland)	6.77	3.10	36.71	33.81	69.81	23.29
Viandra (USSR)	9.07	3.85	136.00	53.60	113.43	34.87
Cirolana (England)	2.74	2.46	82.10	29.57	26.94	13.30
Johan Hjort (Norway)	5.69	0.65	13.15	3.95	33.81	5.69
Skagerak (Sweden)	11.68	5.36	12.74	0.28	95.54	8.03
Thetis (Sweden)	13.44	3.76	3.62	0.32	427.38	23.53
Vaida (USSR)	8.97	18.30	107.85	29.57	105.70	23.78
Aliot (USSR)	-	-	92.69	24.28	82.93	26.39
Feiebas (Norway)	1.56	2.63	131.95	51.46	606.89	84.63
Antares (USSR)	10.47	2.90	95.54	61.80	24.79	21.65

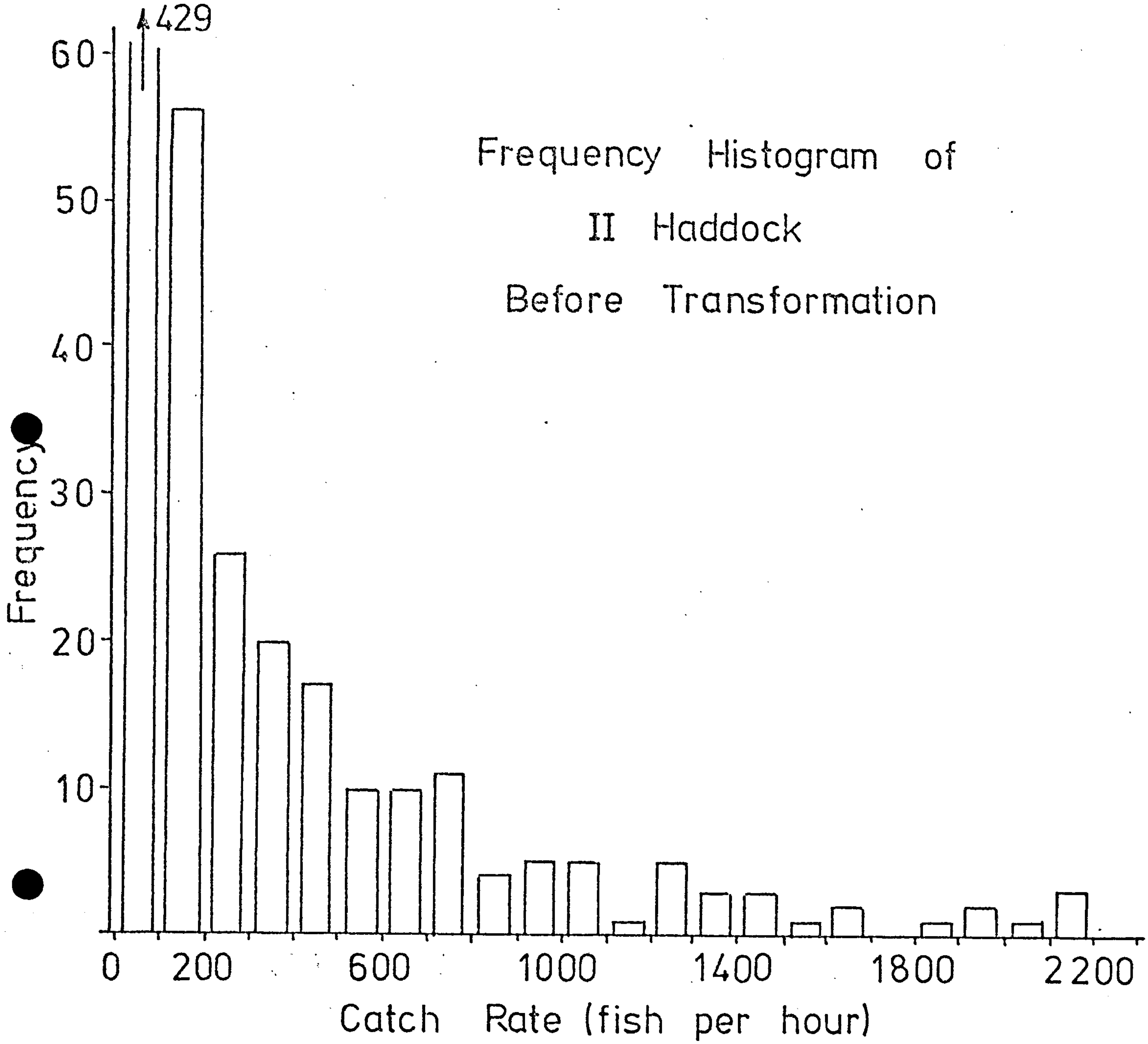

429
60
50
40
30
20
10
Frequency
0
200
600
1000
1400
1800
2200
Catch Rate (fish per hour)
Frequency Histogram of
II Haddock
Before Transformation

FIGURE 2.

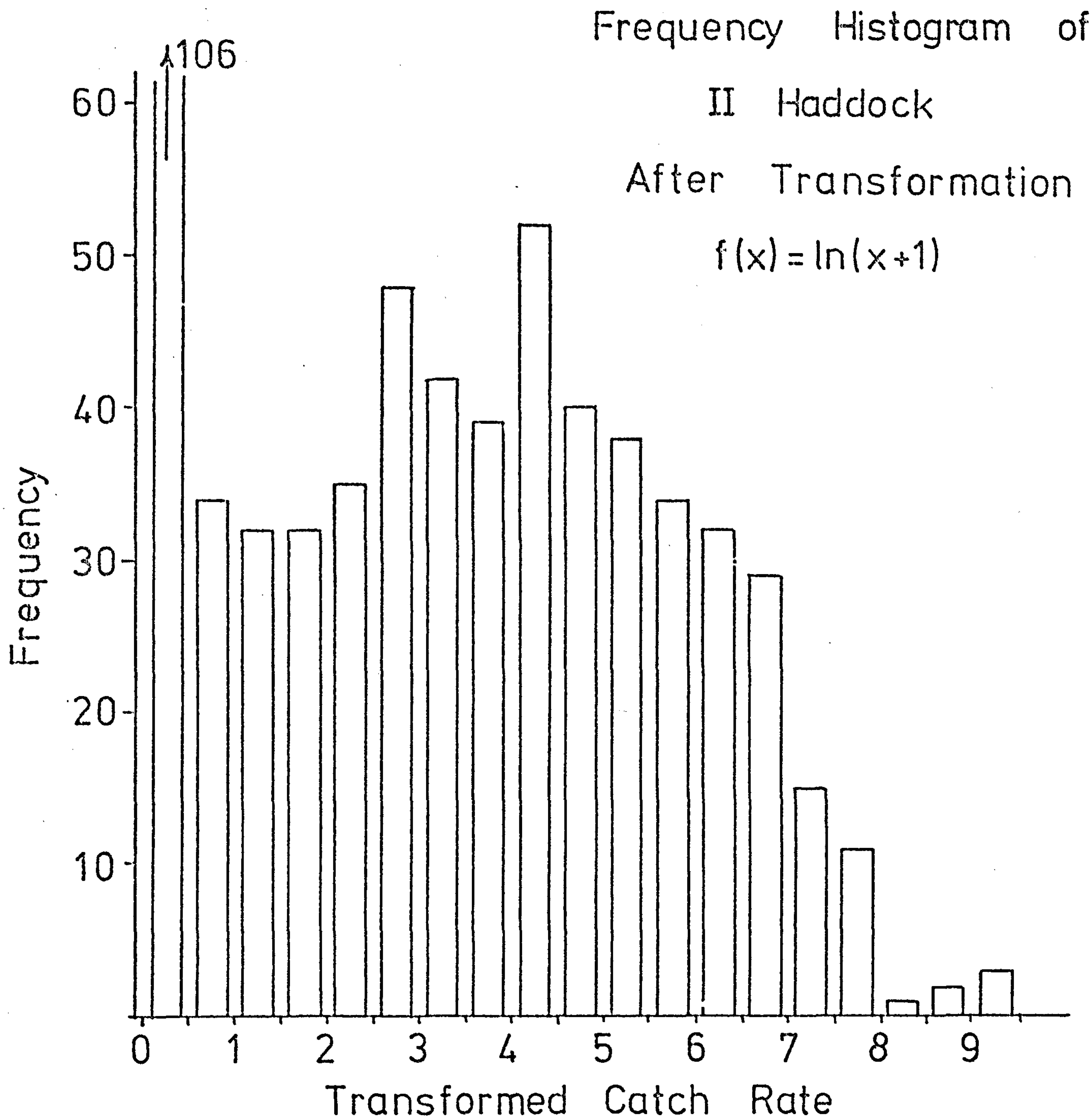
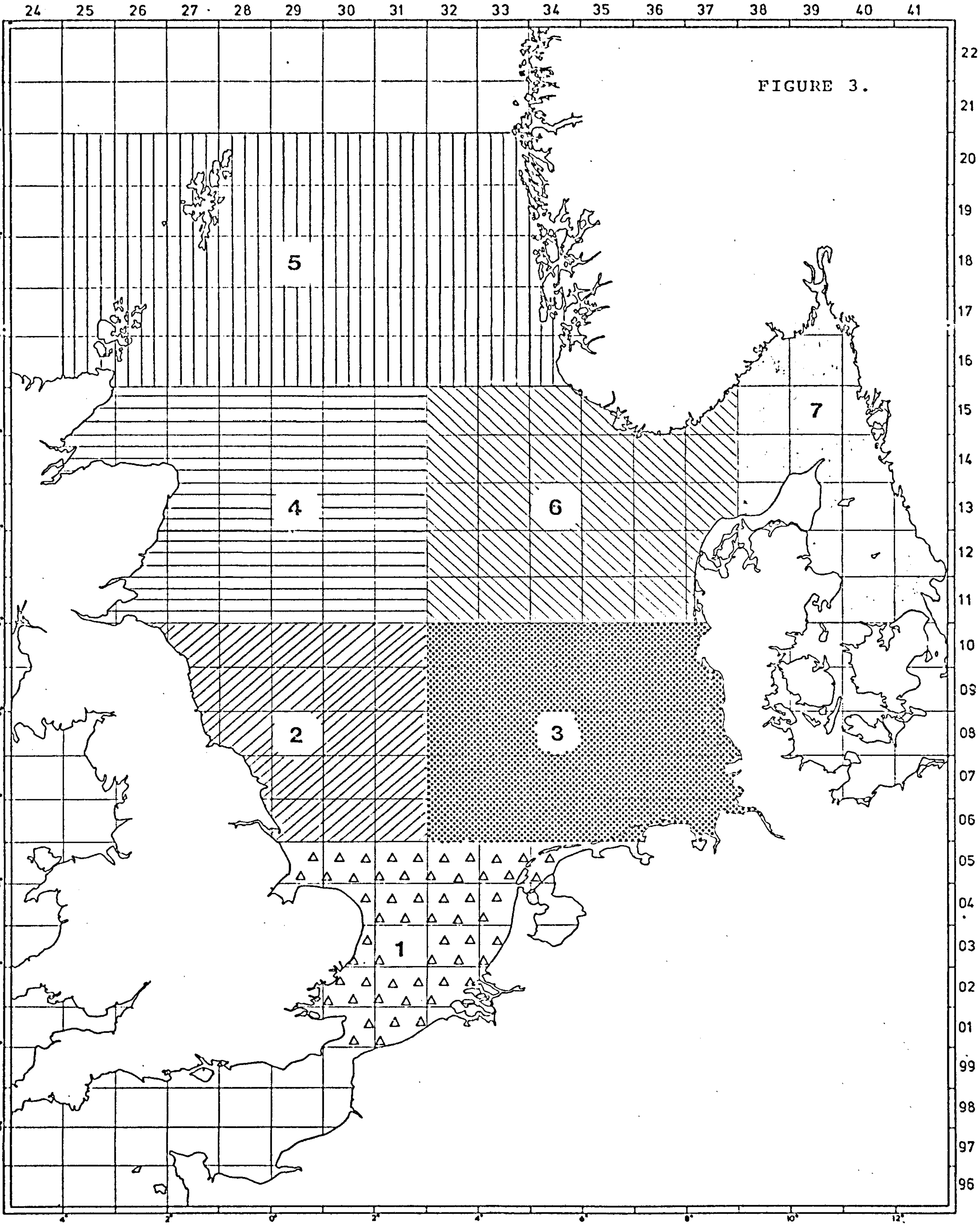

Frequency Histogram of
II Haddock
After Transformation
f(x) = ln(x+1)
106
60
50
40
30
20
10
0
1
2
3
4
5
6
7
8
9
Frequency
Transformed Catch Rate

FIGURE 3.



24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41
22 21 20 19 18 17 16 15 14 13 12 11 10 09 08 07 06 05 04 03 02 01 99 98 97 96
61° 60° 59° 58° 57° 56° 55° 54° 53° 52° 51° 50°
4° 2° 0° 2° 4° 6° 8° 10° 12°
1
2
3
4
5
6
7

FIGURE 4.

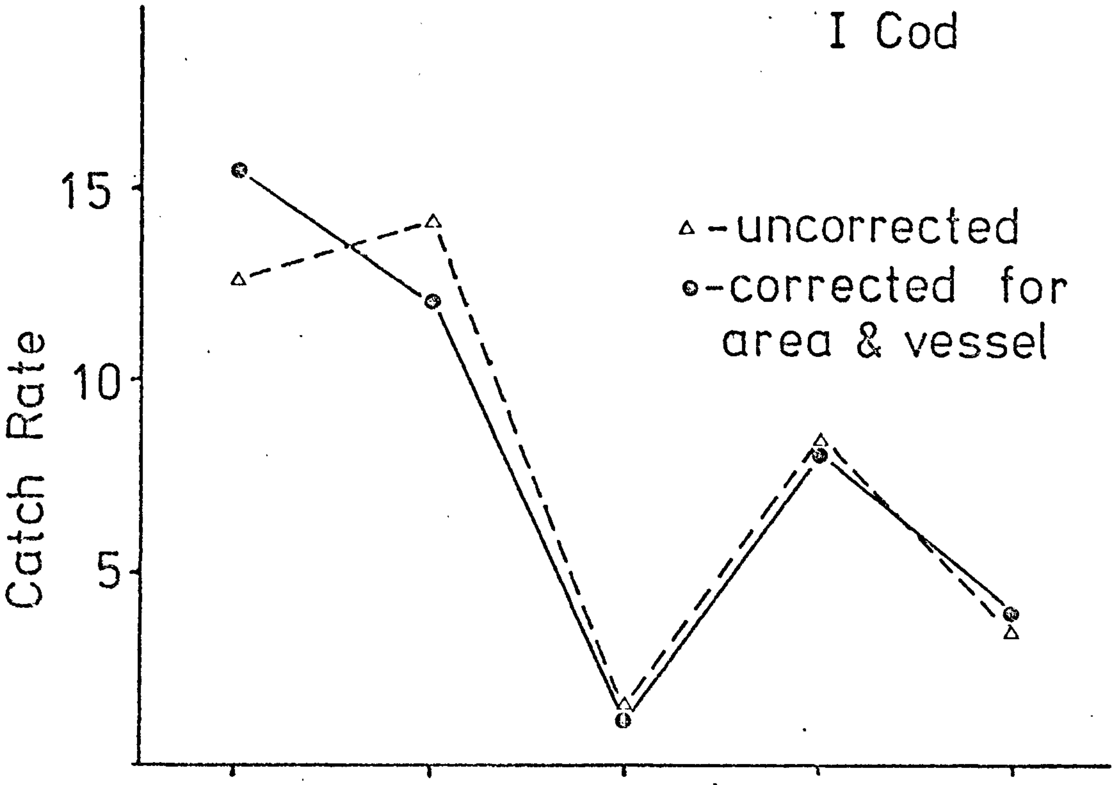


FIGURE 5.

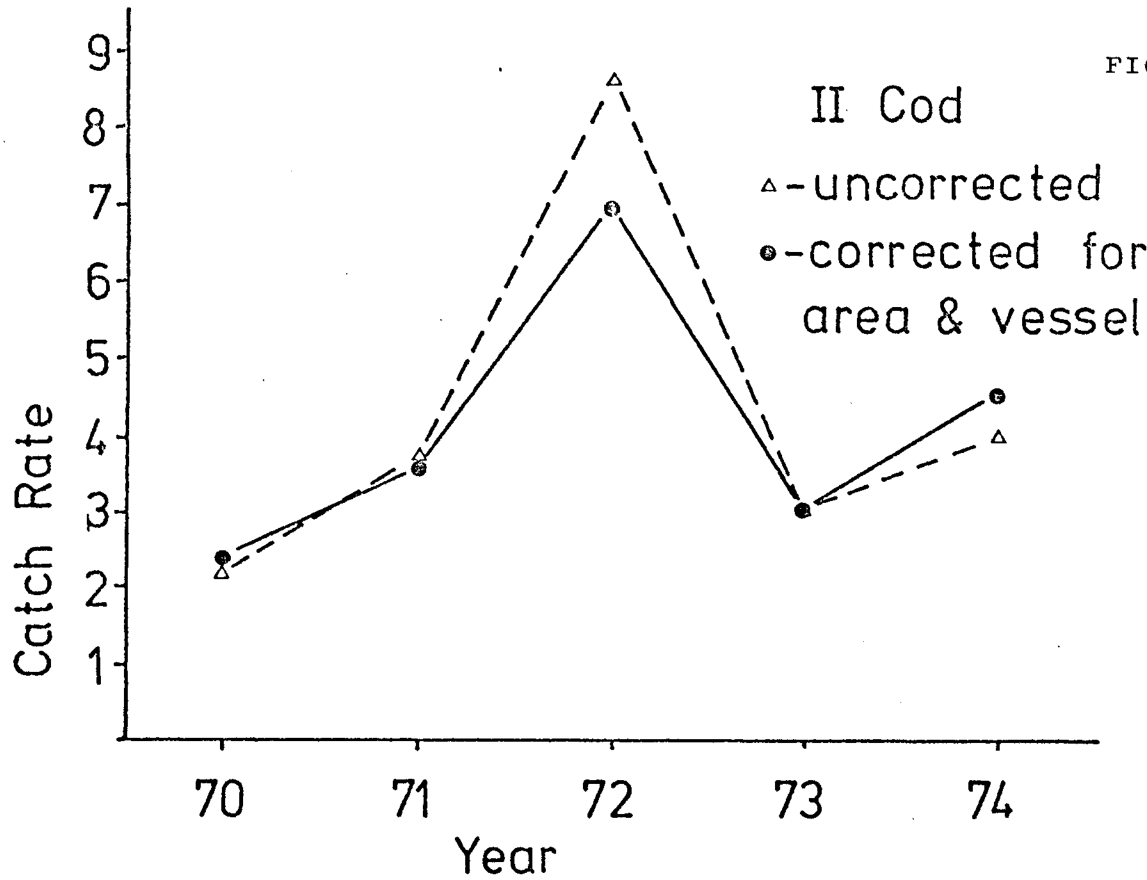
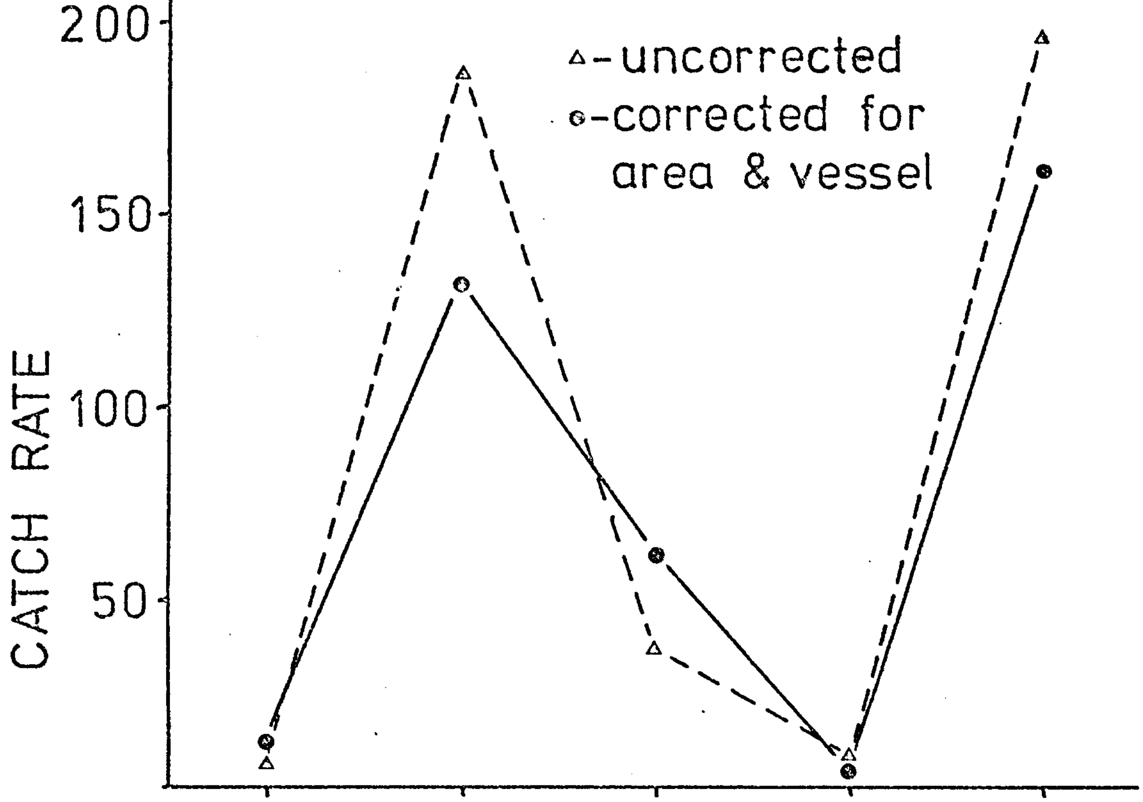


FIGURE 6.



II HADDOCK - FIGURE 7.

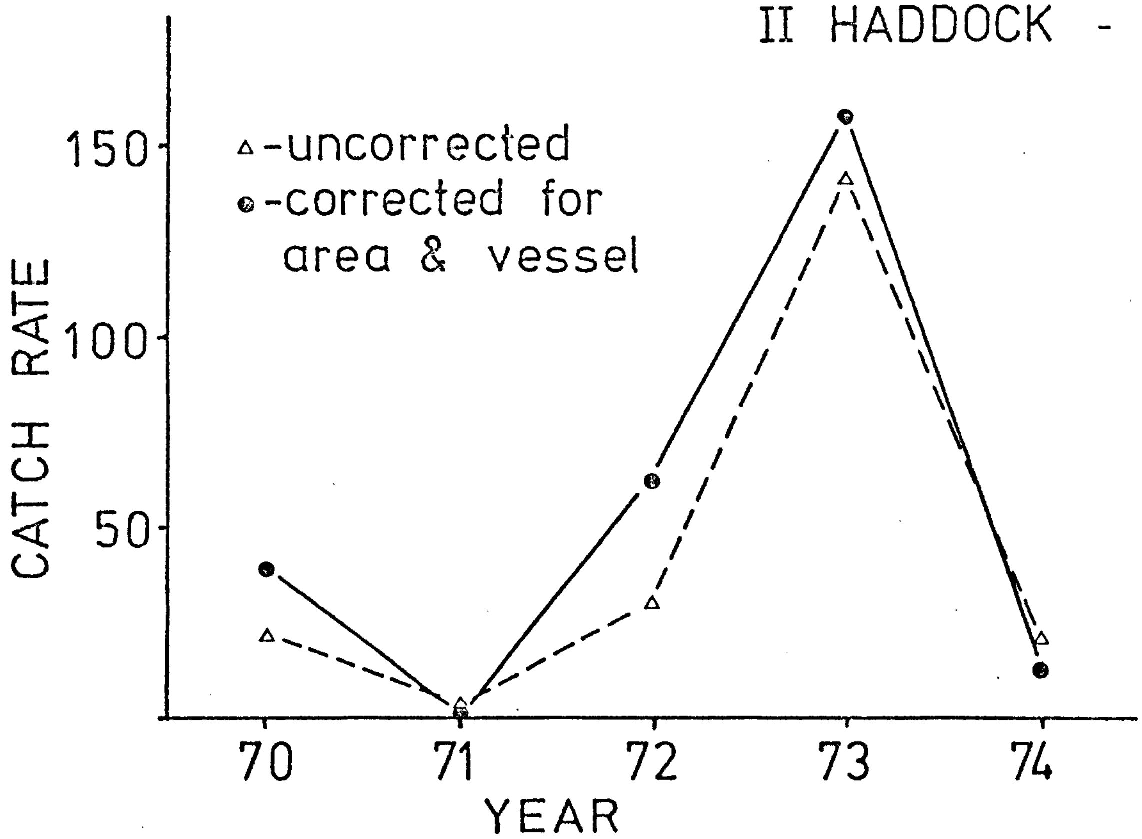


FIGURE 8.

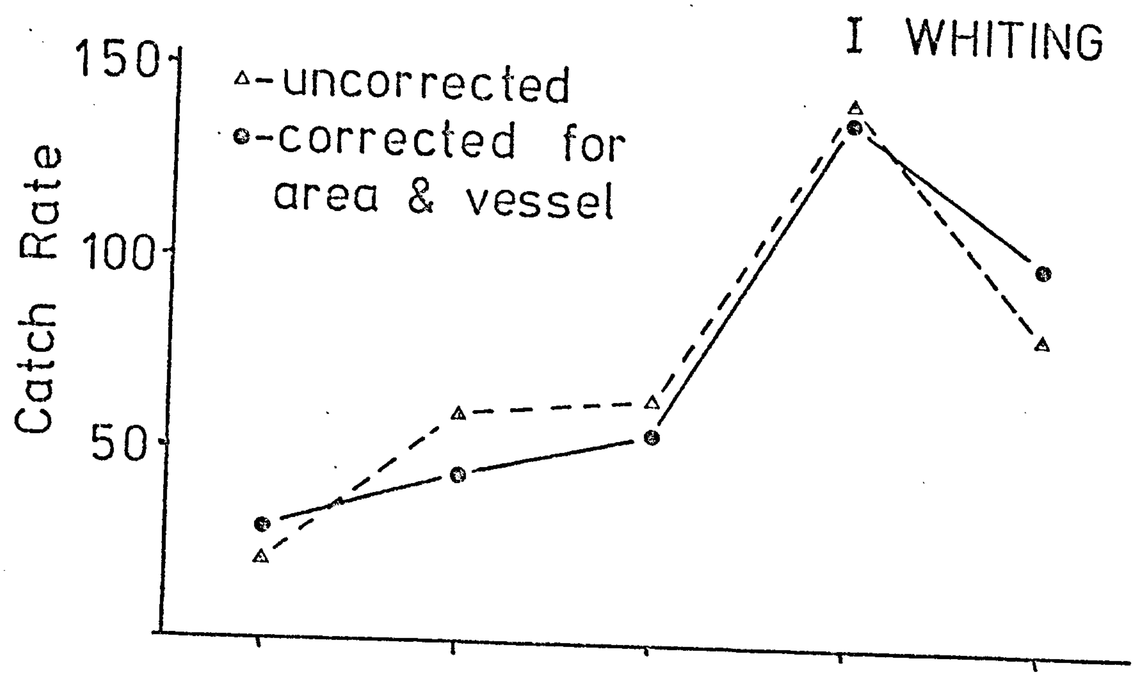


FIGURE 9.