SCRS/1996/094 Col.Vol.Sci.Pap. ICCAT, 47 : 255-265 (1998)

STANDARDIZATION OF CPUE FOR WHITE MARLIN CAUGHT BY JAPANESE LONGLINE FISHERY IN THE ATLANTIC

Yuji Uozumi[1]

SUMMARY

The white marlin CPUE of the Japanese longline fishery from 1960 to 1995 was standardized by Generalized Linear Model (GLM) for the Atlantic Ocean. Two kinds of models with different error structures were used (log-normal and Poisson distribution). The main effects of year, area, season, and gear configuration were incorporated into the models. The standardization was carried out under two hypotheses on the stock structure, one total Atlantic stock or two separate (north and south) Atlantic stocks. There was no large difference in the general trend of standardized CPUEs between the two models. The CPUEs increased rapidly in the 1960s and then continued to decrease, with the rate of decrease becoming smaller in the 1975-1995 period than in the period of 1960-1975.

RÉSUMÉ

La CPUE des makaires blancs de l'Atlantique capturés par les palangriers japonais entre 1960 et 1995 a été standardisée avec le Modèle Linéaire Généralisé (GLM). Deux types de modèles avec différentes structures de l'erreur (lognormale et Poisson) ont été appliqués. Les effets zone, saison et engin ont été pris en compte dans les modèles. La standardisation a été menée dans deux hypothèses de structure du stock : 1) Atlantique entier et 2) Atlantique Nord et Atlantique Sud. La tendance générale de la CPUE est très semblable dans les deux modèles. La CPUE indique un déclin rapide dans les années soixante, suivi d'un déclin plus progressif. Le rythme de diminution est plus faible entre 1975 et 1995 qu'entre 1960 et 1975.

RESUMEN

Se estandarizó la CPUE de aguja blanca de la pesquería de palangre de Japón en el Océano Atlántico desde 1960 a 1995 mediante el Modelo Lineal Generalizado. Se utilizaron dos clases de modelos con diferentes estructuras de error (distribución logarítmica normal y de Poisson). Se incorporaron a los modelos los efectos de área, temporada y configuración del arte. La estandarización se llevó a cabo bajo las dos hipótesis de estructura de stock, uno o dos stocks (norte y sur) en el Atlántico. No hay grandes diferencias entre la tendencia general de la CPUE estandarizada de los dos modelos. La CPUE se incrementó rápidamente en los años 60, y después continuó disminuyendo. La tasa de disminución se hizo menor en el período 1975-95 que en el período 1960-75.

1. INTRODUCTION

The CPUEs of longline are one of the valuable sources of information on the historical changes in stock abundance for the size class vulnerable to longliners. There are, however, many factors which seem to affect CPUE such as environmental factors (season, water temperature, salinity, depth, current, and moon phase), biological factors (maturation, migration, and food availability), and operational factors (gear configuration, soaking time, target species, regulation, kind of bait, and hook size). Therefore, standardization becomes essential for the purposes of monitoring the stock abundance using the CPUE. However, the data availability on such factors is generally very limited and it is possible to incorporate some of these factors into the analysis in the present study.

A multiplicative model has been used in the standardization with the assumption that the error structure has a log-normal distribution (Nakano *et al.*, 1994). One problem is the treatment of zero CPUE in log-transformation. To overcome this problem, an arbitrary constant is added, such as "1" to the zero CPUE. A possible better way to deal with the zero CPUE has been attempted in some recent studies (Miyabe, 1995; Nishida *et al.*,, 1994) using a catch model with Poisson error structure. In the present study these two kinds of model were used for the standardization of CPUE of white marlin data.

[1] National Research Institute of Far Seas Fisheries,Fisheries Agency of Japan, 5-7-1 Orido, Shimizu, 424 Japan

2. DATA AND METHODS

2.1. Basic data

The basic data for this study were obtained from the Japanese longline fishery statistics compiled at the National Research Institute of Far Seas Fisheries for 1956-1995. There are two databases. One contains catch (number) and effort (number of hooks) statistics from 1956 to 1995, aggregated by month, and also five-degree area. This is the TASK II database submitted to ICCAT. The other contains sample statistics, aggregated by the same strata as in the former database, with an additional item of the number of branch lines per basket, which approximates the gear configuration for regular and deep longlines. These sample statistics have been collected since 1975. In the present study, the former database is used from 1960 to 1975, because the distribution and number of data were insufficient prior to 1960. The latter database is used from 1975 to 1995, though the data for 1995 are preliminary. Fishing operations within the 200 mile EEZ area were excluded in the present analysis to avoid any biases introduced by regulations and to keep consistency of the data. Observations with less than 3,000 hooks were also excluded from the analysis.

2.2. Standardization

Year, season, area, and gear configuration were incorporated as the main effects. The gear configuration is only used in the latter period from 1975-95. Quarter-of-the-year was selected as season. Based on the distribution of efforts and CPUE of white marlin, 10 and 12 subareas were selected for the two periods 1960-75 and 1975-95 as shown in Figure 1. With regard to the gear configuration, 3 to 20 hooks between floats were observed in the statistics used. This range was arbitrarily categorized into 4 levels (4-8, 9-11, 12-15, and 16-20 hooks between floats).

The fitting for the CPUE and catch models with log-normal and Poisson distributions was done using GLM and GENMOD procedures of SAS/STAT statistical package (Ver. 6.11) respectively.

The multiplicative CPUE model with log-normal distribution is :

$$\log(CPUE_{ijkl} + \text{Constant}) = f\hat{E} + YR_i + QT_j + Area_k + Bran_l + \text{Interactions} + e_{ijkl} \quad (1)$$

where log : natural logarithm ; $CPUE_{ijkl}$: nominal CPUE (catch in number per 1000 hooks, in year i, quarter j, subarea k, and effect of gear l) ; $f\hat{E}$: overall mean ; YR_i : effect of year i; QT_j : effect of quarter j ; $Area_k$: effect of subarea k ; $Bran_l$: effect of gear l ; Interactions : two-way interactions ; e_{ijkl} : normal error term.

To determine the suitable constant in this model, several distribution patterns of log(CPUE) were observed with different constants including 0.1 x mean CPUE, which was recommended by the bluefin tuna species working group (ICCAT, 1997).

The catch model with Poisson distribution is:

$$E(C) = H \exp(f\hat{E} + YR_i + QT_j + Area_k + Bran_l + \text{Interactions}) \quad (2)$$

where E(C) : expectation of catch in number which belongs to Poisson distribution ; H : number of hooks.

In the GENMOD analysis, over-dispersion was observed, so the DSCALE option, which estimates a dispersion parameter, was included. All statistics were appropriately adjusted according to the dispersion parameter.

The standardization is carried out under the two hypotheses of stock structure, one total Atlantic stock and separate north and south Atlantic stocks, as in Nakano *et al.* (1994). The number of observations by year for each stock hypothesis is shown in Table 1. All combinations of two-way interactions were tested for the selection of the best CPUE model, except for the two-way interactions of year and gear configuration (YR x Bran), due to the lack of sufficient data combinations for this two-way interaction. The best CPUE model for each period under each hypothesis was selected using Akaike's Information Criterion (AIC):

$$AIC = X \log(MSE) + 2Y \quad (3)$$

where X is the number of observation, MSE is mean square error and Y is number of parameters. The best CPUE model was also applied arbitrarily to the catch model.

3. RESULTS AND DISCUSSIONS

Figures 2 and 3 show the distribution patterns of log transformed CPUE for white marlin with various constants for the periods from 1960 to 1975 and 1975 to 1995, respectively. These Figures show that distributions of CPUE with the constant larger than 0.1 x mean CPUE are skewed. However, these results imply that a constant smaller than 0.1 x mean is suitable for log transformation. In the present study, the value of 0.001 was applied for both periods.

The AICs obtained for CPUE models are listed in Table 2 for both periods under the two stock structure hypotheses. For the period 1960 to 1975, the final model for the total Atlantic stock included all main effects (year, area, and quarter) and two two-way interactions (area x quarter and year x quarter). The best models for north and south Atlantic stocks are the same and include all main effects and one two-way interaction (area x quarter). For the period 1975 to 1995, the final model for the total Atlantic stock included all main effects (year, area, quarter, and bran) and two two-way interactions (area x quarter and quarter x bran). The best model for the north Atlantic stock is the same as the one for the total Atlantic. The best model for the south Atlantic stock includes all main effects and one two-way interaction (area x quarter).

The result of the ANOVA for the CPUE model using the GLM procedure and goodness of fit for the catch model using the GENMOD procedure are shown in Tables 3 and 4. The F-values in Table 3 and scaled Pearson$f\hat{O}^2$ show the models are highly significant. Figure 4 shows the standardized residuals of the CPUE model for each period under each stock hypothesis. The distributions of the standardized residuals were not far from normal distribution, though skewed to the negative side under each stock hypothesis in the period from 1960 to 1975. In the period 1975 to 1995 the distribution was skewed to the positive side.

The standardized CPUEs, which were expressed relative to 1.00 in 1975, are shown in Figures 5-8. Although there is no statistical tool to compare the goodness of fit between the CPUE and catch models, especially in the case considering the over-dispersion (ICCAT, 1997), the general trend of CPUE in each case is similar in the two models. Figure 9 shows the relationship in trend of standardized CPUE between the two models. The correlation coefficient is 0.96 for the total Atlantic, and 0.93 both for the north and south Atlantic stocks. The CPUEs in 1960 and 1961 are not included due to their very large values. The relationship of the relative CPUEs between the two models is linear and explained well by the linear regression equation (dotted line) shown in Figure 9.

The standardized CPUE increased rapidly in the early 1960s and then began to decrease gradually up to 1995, with some fluctuations in the total Atlantic stock (Figures 5-8). For the north Atlantic stock, the standardized CPUE increased gradually from 1960 to the mid-1960s and then began to decrease. The CPUE for the south Atlantic stock increased rapidly in the early 1960s and then began to decrease. The rate of decrease became smaller in the period from 1975 to 1995 than that during the period from 1960 to 1975. In the latter period, the rate of decrease for the catch model was smaller than that for the CPUE model, although there was a strong spike in the mid-1970s for the catch model.

The distribution pattern of fishing effort changed significantly in the 1970s due to the change in fishing strategy (Uozumi and Nakano, 1994). As a result, the number of observations also changed (Table 1). There was an insufficient number of observations in the early 1960s for the north Atlantic stock database and in the mid-1970s for the south Atlantic stock database. The insufficient number of the observations may affect the estimation of CPUE. Furthermore, the Japanese longline fishery has historically caught billfishes as by-catch. There is a significant difference in the distribution pattern between CPUE and fishing effort (Uozumi and Nakano, 1994; Uozumi, 1997). This is one of the reasons for the lack of sufficient observations in the main distribution area of white marlin, especially for the period after the 1970s. There is a possibility that this discrepancy may introduce some biases in the CPUE trend. There are many other factors which are not included in the standardization and the low value of R-square suggests these factors may have significant affect.

There are some clear differences in the trend between the two models, such as spikes in the CPUE of the catch model, though the general trend is similar. As mentioned above, there is no objective tool to measure which model is better to explain the CPUE. As the next step, observations on variance vs. mean for CPUE are desirable to examine whether the data fit into any of the error structures and to attempt to determine which is preferable.

4. LITERATURE CITED

ICCAT 1997: Report of the bluefin tuna methodology session. Inter. Comm. Conser. Atl. Tunas, *Coll. Vol. Sci. Pap.* XLVI(1): 187-213.

MIYABE, N. 1995: Updated standardized CPUE of Atlantic bluefin caught by the Japanese longline fishery in the eastern Atlantic and Mediterranean Sea. Inter. Comm. Conser. Atl. Tunas, *Coll. Vol. Sci. Pap.* XLIV(1):335-339.

NAKANO, H. and Y. Uozumi, and M. Honma 1995: The CPUE trend for Atlantic white marlin caught by Japanese longline fishery. Inter. Comm. Conser. Atl. Tunas, *Coll. Vol. Sci. Pap.* XLI, 393-407.

NISHIDA, T., Y. Ishizuka, and H. Kishino 1994: Estimation of catch rates of Japanese longline fisheries for southern bluefin tuna (*Thunnus maccoyii*). The 13th meeting of Australian, Japanese and New Zealand scientists on the southern bluefin tuna. SBFWS/94/11, 30pp.

UOZUMI, Y. and H. Nakano 1992: A historical review of Japanese longline fishery and billfish catches in the Atlantic Ocean. Inter. Comm. Conser. Atl. Tunas, *Coll. Vol. Sci. Pap.* XLI, 233-243.

UOZUMI, Y. 1997: Recent status of the Japanese longliner fishery in the Atlantic Ocean laying stress on billfish catches. Inter. Comm. Conser. Atl. Tunas, *Coll. Vol. Sci. Pap.* This volume.

Table 1. Number of observations by year for each stock hypothesis.

Year	North	South	All
60	74	120	194
61	70	215	285
62	128	241	369
63	197	250	447
64	287	271	558
65	350	331	681
66	228	243	471
67	201	170	371
68	180	167	347
69	170	131	301
70	238	114	352
71	318	143	461
72	167	87	254
73	108	101	209
74	121	35	156
75	172	86	258

Year	North	South	All
75	304	127	431
76	232	52	284
77	164	72	236
78	249	67	316
79	250	145	395
80	355	352	707
81	535	297	832
82	603	441	1044
83	285	296	581
84	336	494	830
85	489	602	1091
86	341	410	751
87	314	308	622
88	373	771	1144
89	753	937	1690
90	514	1009	1523
91	474	902	1376
92	458	712	1170
93	461	917	1378
94	484	1060	1544
95	376	633	1009

Table 2. Estimates of AIC for each CPUE model. The value of AIC underlined shows the best model in each period under each stock hypothesis.

1960-1975

Model	All Atlantic	North	South
YR Area QT	4776.3	2432.0	2306.2
YR Area QT Area*QT	4685.7	**2391.8**	**2287.6**
YR Area QT YR*QT	4818.8	2470.2	2341.1
YR Area QT YR*Area	4912.1	2508.4	2367.2
YR Area QT Area*QT YR*QT	**4677.6**	2449.1	2327.5
YR Area QT Area*QT YR*Area	4818.1	2479.3	2349.8
YR Area QT YR*QT YR*Area	4946.2	2544.5	2404.3
YR Area QT Area*QT YR*QT YR*Area	4862.6	2531.1	2390.4

1975-1995

Model	All Atlantic	North	South
YR QT Area Bran	11746.61	12730.91	10583.73
YR QT Area Bran YR*Area	11902.72	12605.00	10644.04
YR QT Area Bran YR*QT	11732.03	12591.66	10632.37
YR QT Area Bran Area*QT	11490.93	12306.80	**10570.10**
YR QT Area Bran Area*Bran	11743.58	12724.70	10588.66
YR QT Area Bran QT*Bran	11609.93	12470.09	10582.71
YR QT Area Bran YR*Area YR*QT	11920.58	12534.04	10694.12
YR QT Area Bran YR*Area Area*QT	11661.47	12211.63	10635.77
YR QT Area Bran YR*Area Area*Bran	11937.74	12623.71	10629.64
YR QT Area Bran YR*Area QT*Bran	11794.14	12380.69	10647.95
YR QT Area Bran YR*QT Area*QT	11529.57	12260.06	10610.15
YR QT Area Bran YR*QT Area*Bran	11740.55	12595.61	10632.88
YR QT Area Bran YR*QT QT*Bran	11683.95	12473.57	10640.09
YR QT Area Bran Area*QT Area*Bran	11465.90	12272.90	10575.78
YR QT Area Bran Area*QT QT*Bran	**11455.23**	**12057.41**	10575.27
YR QT Area Bran Area*Bran QT*Bran	11618.82	12482.24	10842.11
YR QT Area Bran YR*Area YR*QT Area*QT	11726.11	12207.41	10681.90
YR QT Area Bran YR*Area YR*QT Area*Bran	11954.80	12550.48	10680.07
YR QT Area Bran YR*Area YR*QT QT*Bran	11874.45	12411.47	10704.12
YR QT Area Bran YR*QT Area*QT Area*Bran	11515.56	12238.44	10611.26
YR QT Area Bran YR*QT Area*QT QT*Bran	11533.85	12230.91	10623.82
YR QT Area Bran Yr*QT Area*Bran QT*Bran	11693.11	12483.60	10638.59
YR QT Area Bran Area*QT Area*Bran QT*Bran	11535.44	12195.58	10579.88
YR QT Area Bran YR*Area Area*QT Area*Bran	11676.24	12238.44	10611.26
YR QT Area Bran YR*Area Area*QT QT*Bran	11644.43	12228.91	10623.82
YR QT Area Bran Yr*Area Area*Bran QT*Bran	11829.64	12195.58	10579.88
YR QT Area Bran YR*Area YR*QT Area*QT Area*Bran	11741.42	12208.34	10669.00
YR QT Area Bran YR*QT Area*QT Area*Bran QT*Bran	11523.92	12218.84	10626.03
YR QT Area Bran YR*Area Area*QT Area*Bran QT*Bran	11664.26	12224.85	10629.80
YR QT Area Bran YR*Area YR*QT Area*Bran QT*Bran	11909.98	12431.59	10690.87
YR QT Area Bran YR*Area YR*QT Area*QT QT*Bran	11731.03	12180.69	10695.73
YR QT Area Bran YR*Area YR*QT Area*QT Area*Bran QT*Bran	11751.48	12188.83	10670.67

Table 3. The results of ANOVA for each best CPUE model.

All Atlantic

1960-75 — R-square: 0.36

Source	DF	SS	MS	F Value	Pr > F
Model	99	19531.27	197.29	31.74	0.0001
Error	5614	34897.68	6.22		
Corrected Total	5713	54428.95			

Source	DF	SS	MS	F Value	Pr > F
YR	15	2630.43	175.36	28.21	0.0001
QT	3	435.21	145.07	23.34	0.0001
AREA	9	6713.15	745.91	119.99	0.0001
QT*AREA	27	2092.62	77.50	12.47	0.0001
YR*QT	45	786.83	17.49	2.81	0.0001

1975-95 — R-square: 0.24

Source	DF	SS	MS	F Value	Pr > F
Model	79	23638.66	299.22	75.96	0.0001
Error	18823	74146.38	3.94		
Corrected Total	18902	97785.04			

Source	DF	SS	MS	F Value	Pr > F
YR	20	1252.42	62.62	15.90	0.0001
QT	3	320.22	106.74	27.10	0.0001
AREA	11	12410.52	1128.23	286.42	0.0001
BRAN	3	307.04	102.35	25.98	0.0001
QT*AREA	33	2166.87	65.66	16.67	0.0001
QT*BRAN	9	539.05	59.89	15.21	0.0001

North Atlantic

1960-75 — R-square: 0.33

Source	DF	SS	MS	F Value	Pr > F
Model	38	8713.78	229.31	38.97	0.0001
Error	2970	17475.26	5.88		
Corrected Total	3008	26189.03			

Source	DF	SS	MS	F Value	Pr > F
YR	15	1741.81	116.12	19.74	0.0001
QT	3	937.00	312.33	53.08	0.0001
AREA	5	2369.00	473.80	80.52	0.0001
QT*AREA	15	970.95	64.73	11.00	0.0001

1975-95 — R-square: 0.27

Source	DF	SS	MS	F Value	Pr > F
Model	59	12961.98	219.69	50.65	0.0001
Error	8251	35791.16	4.34		
Corrected Total	8310	48753.14			

Source	DF	SS	MS	F Value	Pr > F
YR	20	589.41	29.47	6.79	0.0001
QT	3	274.83	91.61	21.12	0.0001
AREA	6	4355.68	725.95	167.35	0.0001
BRAN	3	152.66	50.89	11.73	0.0001
QT*AREA	18	1406.25	78.13	18.01	0.0001
QT*BRAN	9	556.37	61.82	14.25	0.0001

South Atlantic

1960-75 — R-square: 0.37

Source	DF	SS	MS	F Value	Pr > F
Model	30	10546.02	351.53	53.14	0.0001
Error	2674	17689.20	6.62		
Corrected Total	2704	28235.22			

Source	DF	SS	MS	F Value	Pr > F
YR	15	1530.30	102.02	15.42	0.0001
QT	3	443.91	147.97	22.37	0.0001
AREA	3	4426.21	1475.40	223.03	0.0001
QT*AREA	9	653.49	72.61	10.98	0.0001

1975-95 — R-square: 0.23

Source	DF	SS	MS	F Value	Pr > F
Model	42	11264.55	268.20	75.12	0.0001
Error	10549	37663.72	3.57		
Corrected Total	10591	48928.27			

Source	DF	SS	MS	F Value	Pr > F
YR	20	872.91	43.65	12.22	0.0001
QT	3	80.12	26.71	7.48	0.0001
AREA	4	6800.21	1700.05	476.16	0.0001
BRAN	3	492.66	164.22	46.00	0.0001
QT*AREA	12	224.82	18.73	5.25	0.0001

Table 4. Goodness of fit for the catch model used the GENMOD procedure in SAS/STAT.

All Atlantic

1960-75

Criterion	DF	Value	Value/DF
Deviance	5614	495058.92	88.18
Scaled Deviance	5614	5614.00	1.00
Pearson chi-square	5614	787479.94	140.27
Scaled Pearson X2	5614	8930.07	1.59
Log Likelihood	.	30151.29	.

1975-95

Criterion	DF	Value	Value/DF
Deviance	1.90E+04	85532.64	4.54
Scaled Deviance	1.90E+04	18823.00	1.00
Pearson chi-square	1.90E+04	230983.32	12.27
Scaled Pearson X2	1.90E+04	50832.04	2.70
Log Likelihood	.	14152.32	.

North Atlantic

Criterion	DF	Value	Value/DF
Deviance	2970	171678.60	57.80
Scaled Deviance	2970	2970.00	1.00
Pearson chi-square	2970	279093.38	93.97
Scaled Pearson X2	2970	4828.25	1.63
Log Likelihood	.	15113.12	.

Criterion	DF	Value	Value/DF
Deviance	8251	48883.41	5.92
Scaled Deviance	8251	8251.00	1.00
Pearson chi-square	8251	112086.86	13.58
Scaled Pearson X2	8251	18919.07	2.29
Log Likelihood	.	8277.55	.

South Atlantic

Criterion	DF	Value	Value/DF
Deviance	2674	328028.23	122.67
Scaled Deviance	2674	2674.00	1.00
Pearson chi-square	2674	561292.74	209.91
Scaled Pearson X2	2674	4575.51	1.71
Log Likelihood	.	14533.75	.

Criterion	DF	Value	Value/DF
Deviance	1.10E+04	35168.69	3.33
Scaled Deviance	1.10E+04	10564.00	1.00
Pearson chi-square	1.10E+04	113146.13	10.71
Scaled Pearson X2	1.10E+04	33986.93	3.22
Log Likelihood	.	4808.13	.

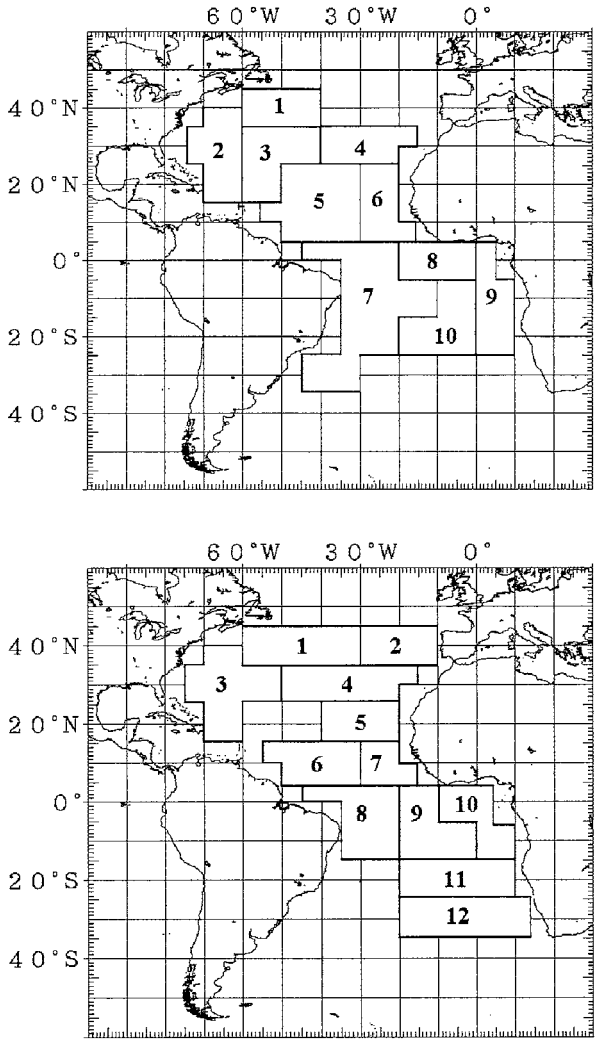

Figure 1. **Subareas used in the standardization of CPUE for white marlin. Top: period of 1960 to 1975, bottom: 1975 to 1995.**

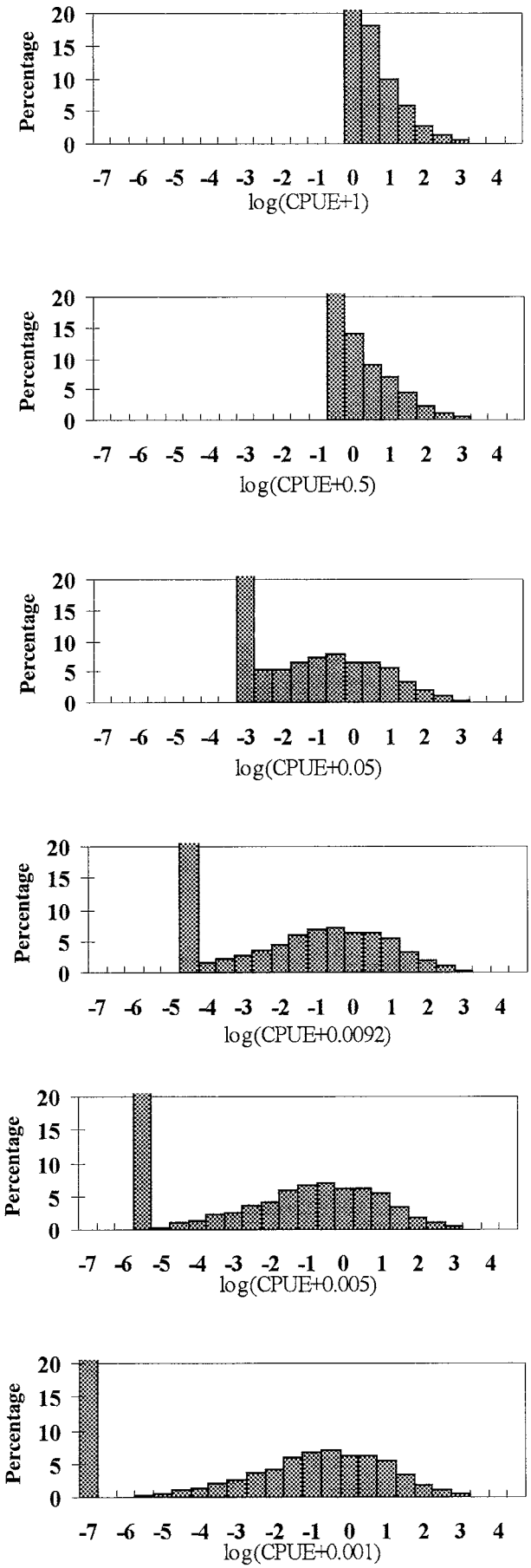

Figure 2. **The distribution pattern of log transformed CPUE of white marlin with various constants for the total Atlantic from 1960 to 1975. The value of 0.0092 is 10% of the overall CPUE.**

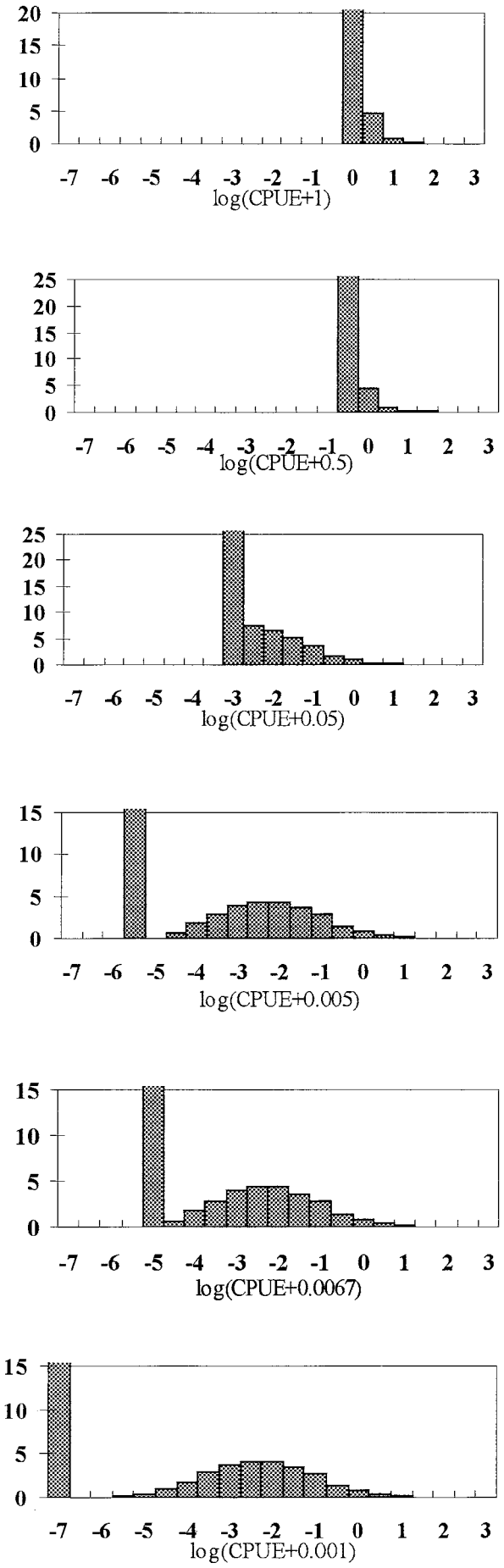

Figure 3. The distribution pattern of log transformed CPUE of white marlin with various constants for the total Atlantic from 1975 to 1995. The value of 0.0067 is 10% of the overall CPUE.

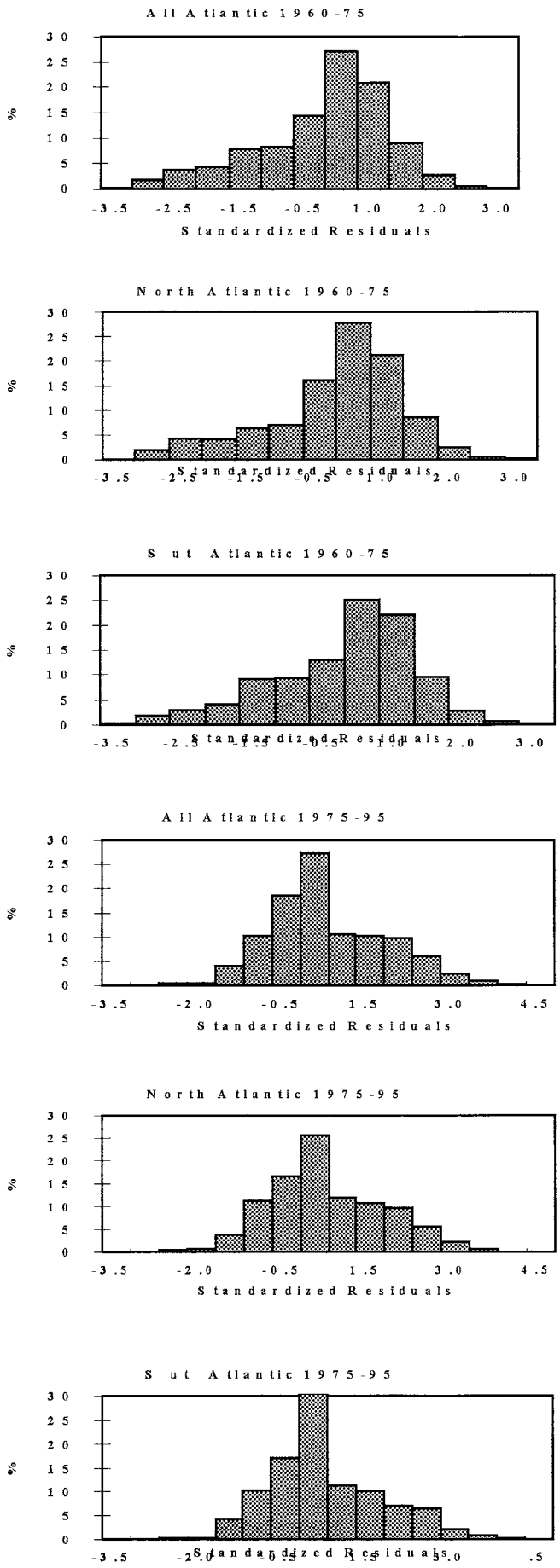

Figure 4. The distribution of standardized residuals of CPUE model in each period under each stock hypothesis.

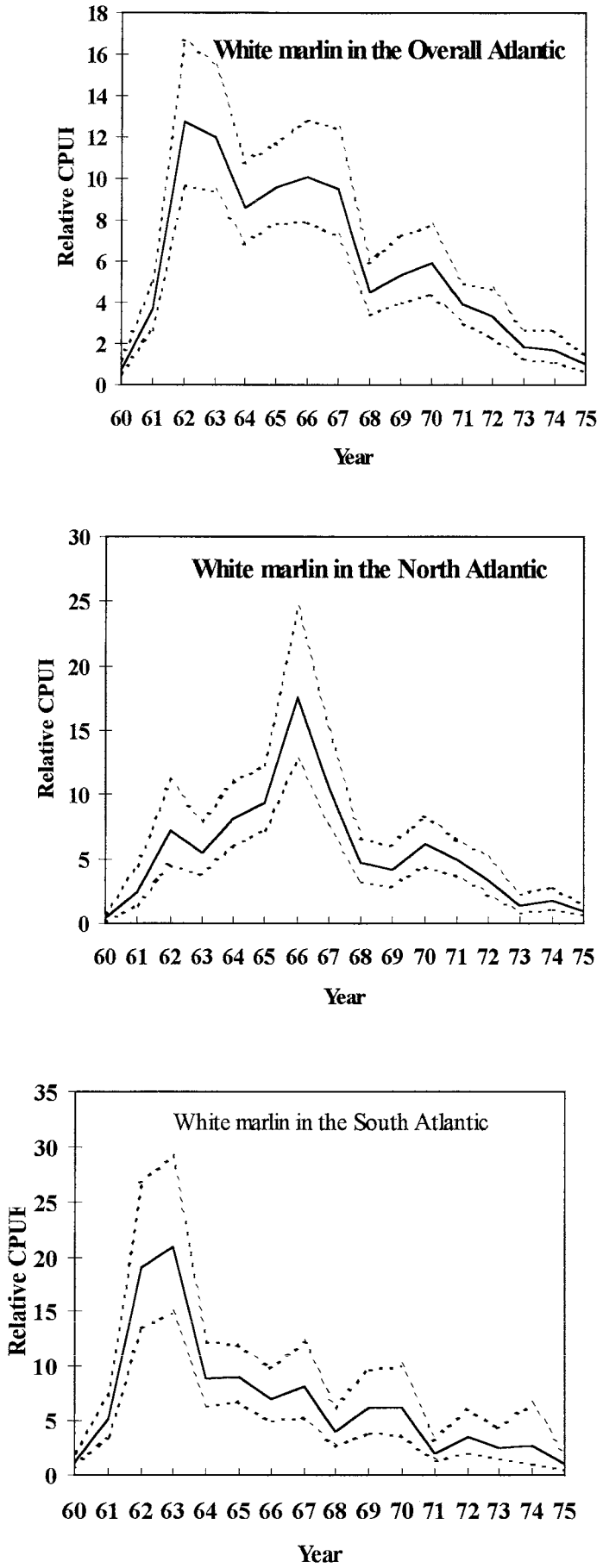

Figure 5. Scaled standardized CPUE of white marlin by CPUE model for the period from 1960 to 1975. The value of the standardized CPUE in 1975 was adjusted to 1.00. Dotted lines show upper and lower 95% confidence limits.

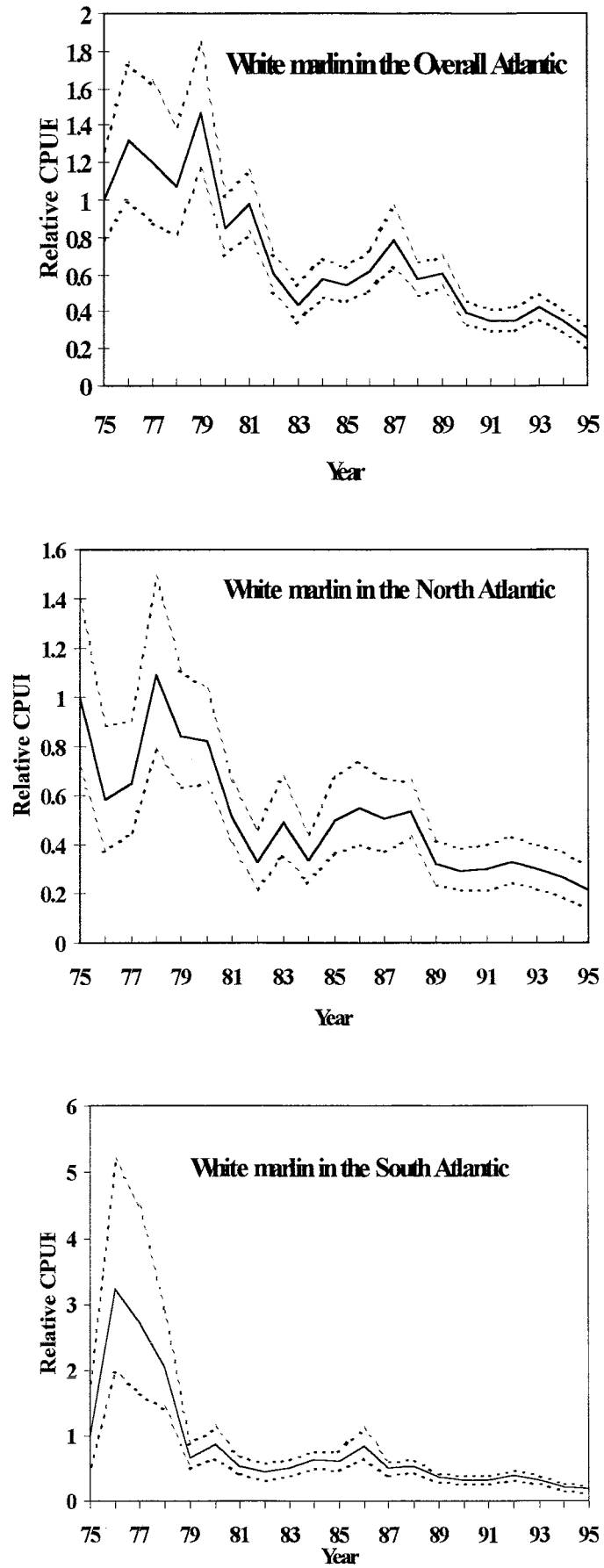

Figure 6. Scaled standardized CPUE of white marlin by CPUE model for the period from 1975 to 1995. The value of the standardized CPUE in 1975 was adjusted to 1.00. Dotted lines show upper and lower 95% confidence limits.

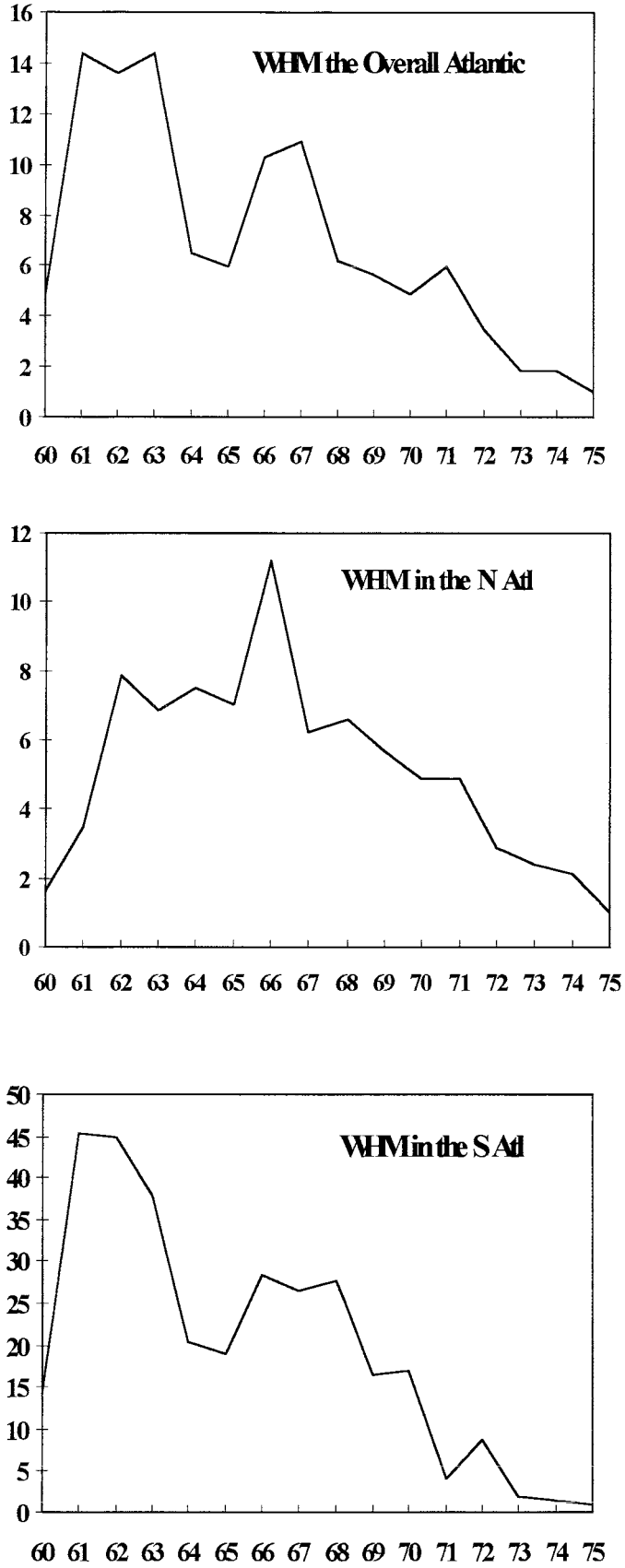

Figure 7. Scaled standardized CPUE of white marlin by catch model for the period from 1960 to 1975. The value of the standardized CPUE in 1975 was adjusted to 1.00.

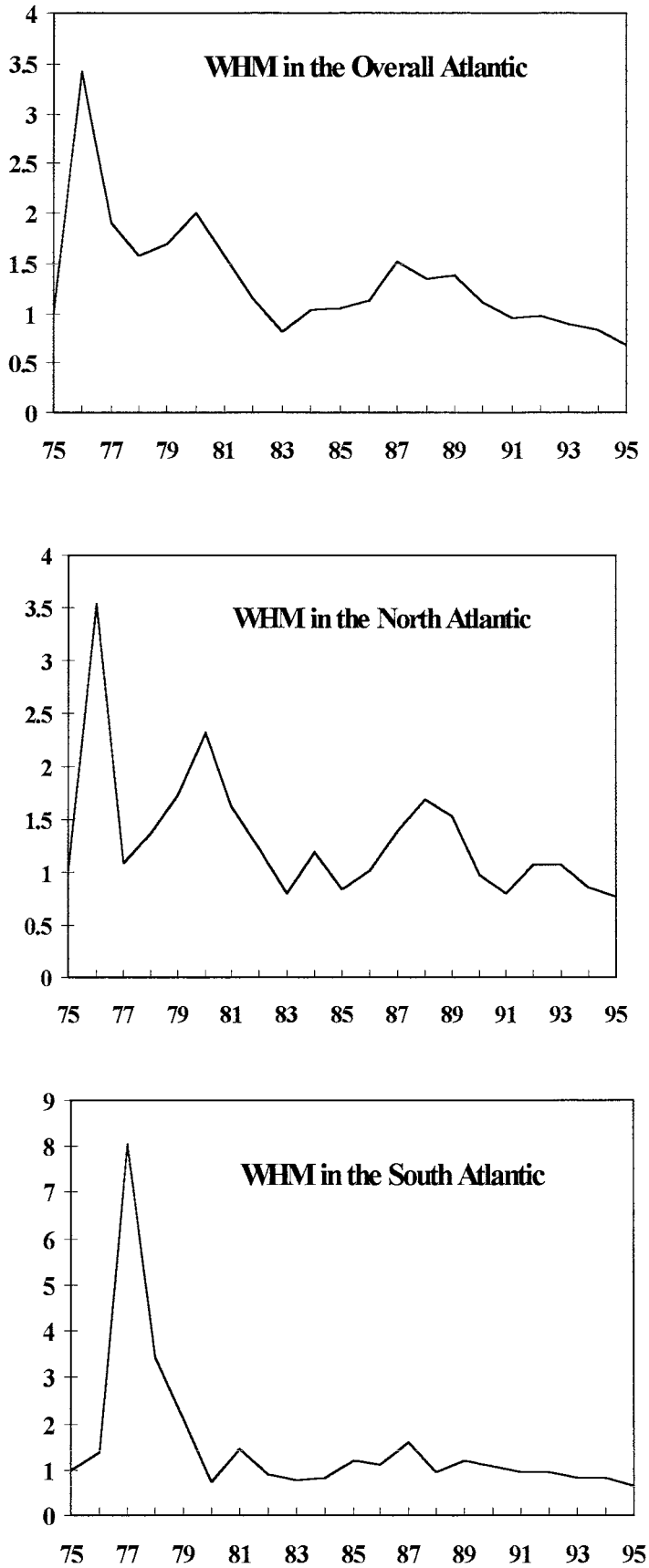

Figure 8. Scaled standardized CPUE of white marlin by catch model for the period from 1975 to 1995. The value of the standardized CPUE in 1975 was adjusted to 1.00.

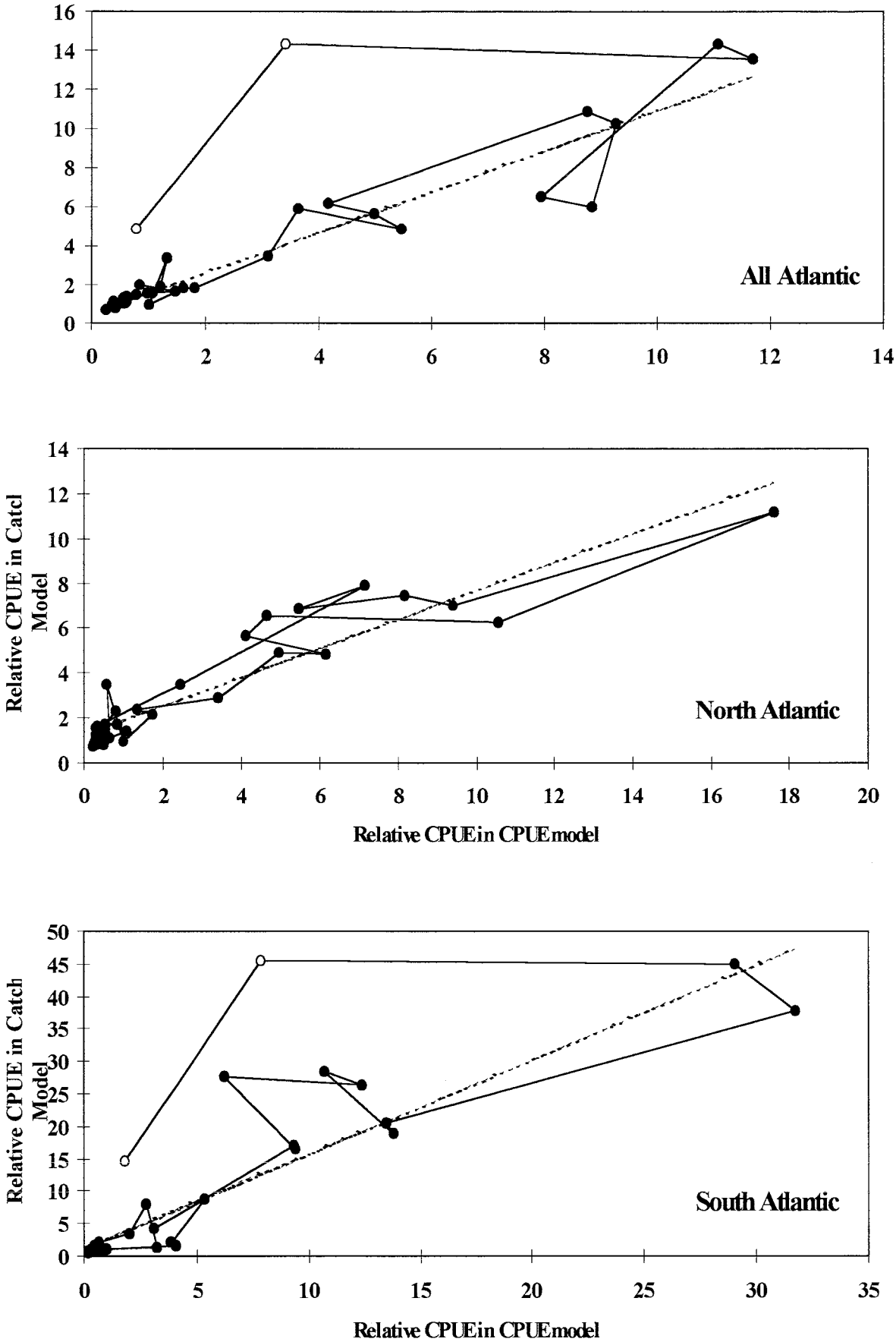

Figure 9. Relationship of scaled standardized CPUE between the two models. Open circles show the CPUEs in 1960 and 1961. Dotted line shows the estimated linear regression line between the two models.