Standardized CPUE for yellowfin tuna (*Thunnus albacares*) of the Japanese longline fishery in the Indian Ocean up to 2003 by Generalized Linear Models (GLM) (1960-2003)

Hiroshi SHONO, Hiroaki OKAMOTO and Tom NISHIDA

National Research Institute of Far Seas Fisheries
5-7-1 Orido Shimizu-ku Shizuoka-shi, 424-8633, Japan

1. Introduction

There are several studies regarding CPUE standardization of yellowfin tuna in the Indian Ocean using generalized linear models (GLM) (Nishida, 2000: Shono *et al.*, 2002). However, model-based stock assessment about yellowfin tuna has not been done since 2002 because of the lack of the size data (Nishida and Shono, 2002). In this report, we carried out the CPUE standardization of yellowfin tuna in the Indian Ocean using updated Japanese longline data of 2001-2003 by the GLM and estimated CPUE index (i.e. year trend of abundance) as an input for stock assessment model such as an age-specific production model (ASPM).

2. Materials and Methods

Data

The Japanese longline catch and effort statistics from 1975 to 2000, which were aggregated by month, 5x5 square area and the number of hooks between floats (NHF), were used for the analysis. Similar data to 1974 were used, although it did not include the information on NHF. Then, we assume that the NHF is 5. SST (sea surface temperature) and MLD (mixed layer depth) data were downloaded from NEAR-GOOS Regional Real Time Data Base of Japan Meteorological Agency and JEDEC (Joint Environmental Data Analysis Center) website of Scripps Institution of Oceanography, respectively. The procedure of these data processing is described in Okamoto *et al.*(2001). (JMA web site: http://goos.kishou.go.jp/rrtdb/database.html and JEDEC web site: http://jedac.ucsd.edu/DATA_IMAGES/index.html)

Model configuration

In order to standardize CPUE of yellowfin, we used the generalized linear models in this analysis. We used the CPUE with log-normal error and the calculation was performed through GLM procedure of SAS/STAT package (Version 9.1.3). At first, the following form was assumed as a full model.

$$\log(\text{CPUE+constant}) = \text{INTERCEPT} + \text{YEAR} + \text{MONTH} + \text{AREA} + \text{NHF} + \text{SST} + \text{MLD} + \text{INTERACTIONS} + \text{ERROR}, \quad \text{ERROR} \sim \text{N}(0, \sigma^2) \qquad (1)$$

where
log: national logarithm,
CPUE: nominal CPUE (number of yellowfin catch per 1000 hooks),
INTERCEPT: intercept,
YEAR: effect of year,
MONTH: effect of month,
AREA: effect of area,
NHF: effect of NHF,
SST: effect of sea surface temperature,
MLD: effect of mixed layer depth,
INTERACTIONS: two-way interactions.

Precisely, formula (1) is written in formula (2). In order to overcome the problem of zero catch, 0.1 was uniformly added to each value of nominal CPUE as the constant term. The small constant value was selected to decrease the bias of confidence intervals in such calculation. The following two-way interactions were used as a full model. Other interactions could not be included into this full model because of missing data.

- *full model* -

$$\begin{aligned}\log(CPUE_{ijkl} + 0.1) &= INTRECEPT + YEAR_i + MONTH_j + AREA_k + NHF + SST \\ &+ MLD + (YEAR*MONTH)_{ij} + (YEAR*AREA)_{ik} + (MONTH*AREA)_{jk} \\ &+ (YEAR*SST)_i + (YEAR*MLD)_i + (AREA*NHF)_k + (AREA*SST)_k \\ &+ (AREA*MLD)_k + (MONTH*NHF)_j + (MONTH*SST)_j + (MONTH*MLD)_j \\ &+ (SST*MLD) + ERROR_{ijk}, \qquad ERROR_{ijk} \sim N(0, \sigma^2)\end{aligned} \qquad (2)$$

where
i (YEAR): 1960-2003,
j (AREA): 1-5,
k (MONTH): 1-12,
Remark) NHF, SST and MLD were incorporated as a continuous variable.

We used the data from 1960 to 2003 and tried to do three scenarios changed the beginning year of calculation and/or sub-area used (shown in Table 1). We used the agreed area stratification at the 2002 WPTT meeting in Shanghai shown in Figure 1.

In our calculation, YEAR, AREA, MONTH, NHF, SST and MLD were incorporated as the main effect. We also added the effect of SST and MLD to the model as environmental factors that may have affected the yellowfin stock status and recruitment.

Model selection

We performed the model selection using the stepwise F-test based on the value of deviance (Dobson, 1990). As a result of the test all about the path that can be considered, the following model with many explanatory variables was finally selected. (i.e. Only the two-way interaction of SST and MLD was deleted from the full model in all three scenarios.) Significant level was assumed to be one percentage. The results of ANOVA are shown in Table 3-5.

- *final model* -

$$\begin{aligned} E[\log(CPUE_{ijk}+0.1)] &= INTRECEPT + YEAR_i + MONTH_j + AREA_k + NHF + SST \\ &+ MLD + (YEAR*MONTH)_{ij} + (YEAR*AREA)_{ik} + (MONTH*AREA)_{jk} \\ &+ (YEAR*SST)_i + (YEAR*MLD)_i + (AREA*NHF)_k + (AREA*SST)_k \\ &+ (AREA*MLD)_k + (MONTH*NHF)_j + (MONTH*SST)_j + (MONTH*MLD)_j, \end{aligned} \quad (3)$$

CPUE and Abundance Index

CPUE index in year i and in a whole area is estimated by the following equation.

$$CPUE_i = \exp\{INTERCEPT + YEAR_i + \overline{(YEAR*AREA)}_{i\cdot} + \overline{(YEAR*MONTH)}_{i\cdot} + (YEAR*SST)_{i\cdot}*SSTmean + (YEAR*MLD)_{i\cdot}*MLDmean\} - 0.1, (i = 1960, \cdots, 2003) \quad (4)$$

where

$$\overline{(YEAR*AREA)}_{i\cdot} = \frac{1}{N_k}\sum_{k=1}^{N_k}(YEAR*AREA)_{ik},$$

$$\overline{(YEAR*MONTH)}_{i\cdot} = \frac{1}{N_j}\sum_{j=1}^{N_j}(YEAR*MONTH)_{ij},$$

$SSTmean$: sample mean of SST, $MLDmean$: sample mean of MLD

The terms of $CPUE_i$+constant (i.e. 0.1) means the Least Squared Means (LSMEANS) of YEAR effect in GLM procedure of SAS package.

The CPUE index in year i and area k is obtained from the following formula.

$$CPUE_{ik} = \exp\{Intercept + YEAR_i + AREA_k + (YEAR*AREA)_{ik} + \overline{(YEAR*MONTH)}_{i\cdot} + (YEAR*SST)_{i\cdot}*SSTmean + (YEAR*MLD)_{i\cdot}*MLDmean + \overline{(AREA*MONTH)}_{k\cdot} + (AREA*NHF)_{k\cdot}*NHFmean + (AREA*SST)_k*SSTmean + (AREA*MLD)_k*MLDmean\} - 0.1, \quad (i = 1960, \cdots, 2003 : k = 1, \cdots, 5) \quad (5)$$

where

$$\overline{(AREA*MONTH)}_{k\cdot} = \frac{1}{N_j}\sum_{j=1}^{N_j}(AREA*MONTH)_{kj},$$

$NHFmean$: sample mean of NHF etc.

The terms of $CPUE_{ik}$+constant (i.e.0.1) means the Least Squared Means (LSMEANS) of (YEAR*AREA) effect in GLM procedure of SAS package.

Abundance index in i year and in a whole area is estimated by the following equation.

$$Abundance\ Index_i = \sum_{k=1}^{N_k} w_k \cdot CPUE_{ik}, \quad (i = 1960, \cdots, 2003) \quad (6)$$

where

$w_k = \text{(Size of area k)} / \text{(Size of total area)}, \quad \sum_{k=1}^{N_k} w_k = 1$

The relative area size is shown in Table 2.

3. Results and Discussions

Figure 2 shows the comparison of the absolute year trends between standardized and nominal CPUE in the Base-case. Those two trends are considerable similar until 1970s. The trend of standardized CPUE is smoother than that of nominal one since 1980. It is seemed to be natural because CPUE standardization probably corrects the various effects in this analysis.

Figure 3, 4 and 5 show the standardized CPUE trend & its ninety-five percent confidence intervals and corresponding abundance index (that is weighted by relative area-size) in the Base-case, Option-1 and Option-2, respectively. Its confidence intervals are totally narrow, especially after the period of 1980s. In all three scenarios, the difference of the year trends between standardized CPUE and abundance index is not so large.

Figure 6 and 7 show the relative year trends of standardized CPUE (that is not weighted by area size) and abundance index (that is weighted by area size) in all three scenarios (i.e. Base-case, Option-1 and Option-2). These year trends of CPUE and abundance index are rather similar except the period of 1960s. Year trends of Option-2 are corresponding to Taiwanese indices and these values are a little higher than other CPUE series since 1980s. The distribution of overall residual in the final model is shown in Figure 7. Judging from the distributed pattern (i.e. Histogram and QQ-plot) of three scenarios, the log-normal models we used this time seem to be appropriate (Figure 8-10).

The year trends of standardized CPUE/abundance-index seem to be rather stable in the 1980s and 1990s, although small yellowfin catch by purse fishery has been increased rapidly after the middle of 1980s. This implies that standardized CPUE by GLM may not reflect the real stock status for yellowfin tuna in the Indian Ocean, especially after in the middle of 1980s. And also it seems to be needed to monitor the influence of large catch for yellwofin tuna for 2003-2004 and discuss those CPUE trends from the biological point of view.

References

Dobson, A. J. (1990): *An introduction to generalized linear models*. Chapman and Hall. 174pp.

Nishida, T. (2000): Standardization of the Japanese longline catch rates of adult yellowfin tuna (*Thunnus albacares*) in the western Indian Ocean by General Lenear Model (1975-98). IOTC/WPTT/00/10, 7p.

Nishida, T. and Shono, H. (2002): Stock assessment of yellowfin tuna (*Thunnus albacares*) resources in the Indian Ocean by the age structured production model analyses (ASPM). IOTC/WPTT/02/13, 28p.

Okamoto, H., Miyabe, N. and Matsumoto. (2001): GLM analyses for standardization of Japanese longline CPUE for bigeye tuna in the Indian Ocena. IOTC/WPTT/01/21, 38p.

Shono, H., Okamoto, H. and Nishida, T. (2002). Standardized CPUE for yellowfin tuna (Thunnus albacares) resources in the Indian Ocean by Generalized Linear Models (GLM) (1960-2000). IOTC/WPTT/02/12. 12p.

Table 1 Three scenarios we calculated in our preliminary analyses.

Sconario	Year	Area
Base-case	1960-2003	Whole area (1,2,3,4,5)
Option-1	1960-2003	Main ground (2 and 5)
Option-2	1968-2003	Whole area (1,2,3,4,5)

Table 2 Area Index for weighting.

Area	Scaled size of each area	Area index for weighting
1	1,727	0.1129
2	3,118	0.2037
3	1,493	0.0976
4	5,871	0.3836
5	3,095	0.2022
Total	15,304	1.0000

Table 3 ANOVA table in the scenario of our Base-case. ("eda" shows NHF.)

Source	DF	Sum of Square	Mean Square	F Value	Pr > F
Model	881	85328.8304	96.8545	68.83	<.0001
Error	48475	68209.0884	1.4071		
Corrected Total	49356	153537.9187			
R-Square	Coeff Var	Root MSE	logCPUE Mean		
0.555751	143.1681	1.186212	0.828545		
Source	DF	Type III SS	Mean Square	F Value	Pr > F
area	4	943.784489	235.946122	167.68	<.0001
year	43	800.85857	18.624618	13.24	<.0001
month	11	175.203942	15.927631	11.32	<.0001
eda	1	25.330555	25.330555	18	<.0001
sst	1	257.912203	257.912203	183.29	<.0001
mld	1	14.352573	14.352573	10.2	0.0014
area*year	172	2875.434164	16.71764	11.88	<.0001
area*month	44	2417.281608	54.938218	39.04	<.0001
year*month	473	2654.163651	5.61134	3.99	<.0001
sst*year	43	1086.055671	25.257109	17.95	<.0001
mld*year	43	282.136528	6.561315	4.66	<.0001
eda*area	4	176.989056	44.247264	31.45	<.0001
sst*area	4	910.048306	227.512076	161.69	<.0001
mld*area	4	350.552124	87.638031	62.28	<.0001
eda*month	11	377.731867	34.339261	24.4	<.0001
sst*month	11	313.616431	28.510585	20.26	<.0001
mld*month	11	148.444112	13.494919	9.59	<.0001

Table 4 ANOVA table in the scenario of our Option-1. ("eda" shows NHF.)

Source	DF	Sum of Square	Mean Square	F Value	Pr > F
Model	707	15577.46056	22.03318	22.09	<.0001
Error	25010	24944.68613	0.99739		
Corrected Total	25717	40522.14669			
R-Square	Coeff Var	Root MSE	logCPUE Mean		
0.384418	67.35138	0.998693	1.482811		
Source	DF	Type III SS	Mean Square	F Value	Pr > F
area	1	117.097358	117.097358	117.4	<.0001
year	43	144.051473	3.350034	3.36	<.0001
month	11	54.210519	4.928229	4.94	<.0001
eda	1	82.559231	82.559231	82.78	<.0001
sst	1	78.7043	78.7043	78.91	<.0001
mld	1	21.086437	21.086437	21.14	<.0001
area*year	43	480.795687	11.181295	11.21	<.0001
area*month	11	365.039965	33.185451	33.27	<.0001
year*month	473	1954.59368	4.132333	4.14	<.0001
sst*year	43	177.65092	4.131417	4.14	<.0001
mld*year	43	159.356044	3.705955	3.72	<.0001
eda*area	1	20.734581	20.734581	20.79	<.0001
sst*area	1	83.60299	83.60299	83.82	<.0001
mld*area	1	191.750751	191.750751	192.25	<.0001
eda*month	11	133.964438	12.178585	12.21	<.0001
sst*month	11	53.592136	4.872012	4.88	<.0001
mld*month	11	118.689714	10.789974	10.82	<.0001

Table 5 ANOVA table in the scenario of our Option-2. ("eda" shows NHF.)

Source	DF	Sum of Square	Mean Square	F Value	Pr > F
Model	737	69072.3519	93.721	62.77	<.0001
Error	42388	63287.8993	1.4931		
Corrected Total	43125	132360.2511			
R-Square	Coeff Var	Root MSE	logCPUE Mean		
0.521851	184.2882	1.221909	0.663042		
Source	DF	Type III SS	Mean Square	F Value	Pr > F
area	4	750.867007	187.716752	125.73	<.0001
year	35	751.201028	21.462887	14.38	<.0001
month	11	200.34981	18.213619	12.2	<.0001
eda	1	23.092134	23.092134	15.47	<.0001
sst	1	198.699368	198.699368	133.08	<.0001
mld	1	18.550684	18.550684	12.42	0.0004
area*year	140	2646.704627	18.905033	12.66	<.0001
area*month	44	2280.433823	51.828041	34.71	<.0001
year*month	385	2288.076529	5.943056	3.98	<.0001
sst*year	35	1077.112168	30.774633	20.61	<.0001
mld*year	35	277.39719	7.925634	5.31	<.0001
eda*area	4	167.395929	41.848982	28.03	<.0001
sst*area	4	717.971818	179.492954	120.22	<.0001
mld*area	4	331.089212	82.772303	55.44	<.0001
eda*month	11	286.123813	26.011256	17.42	<.0001
sst*month	11	298.159142	27.105377	18.15	<.0001
mld*month	11	133.561718	12.141974	8.13	<.0001

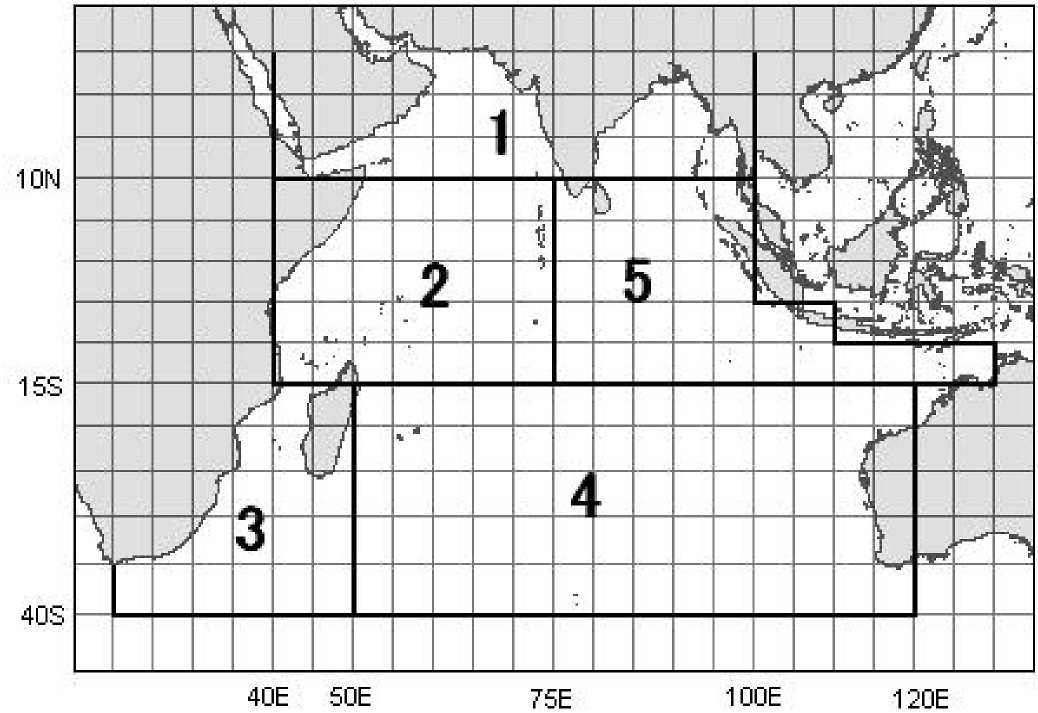


Figure 1 Area stratification agreed in the 2002 WPTT meeting in Shanghai.

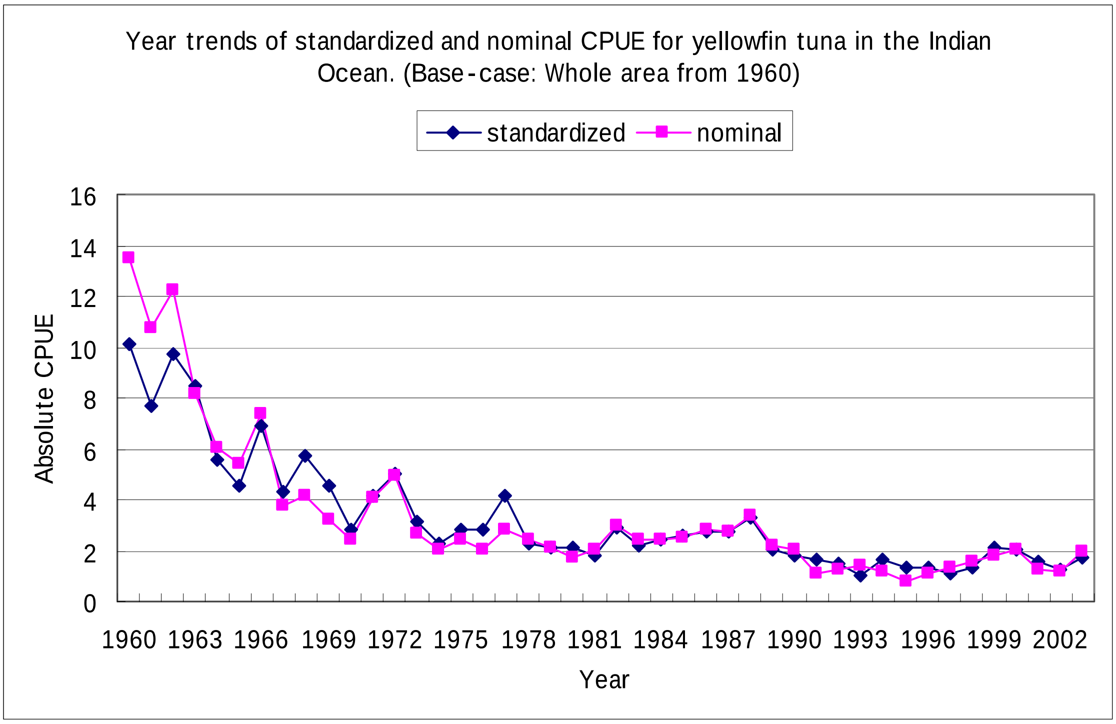
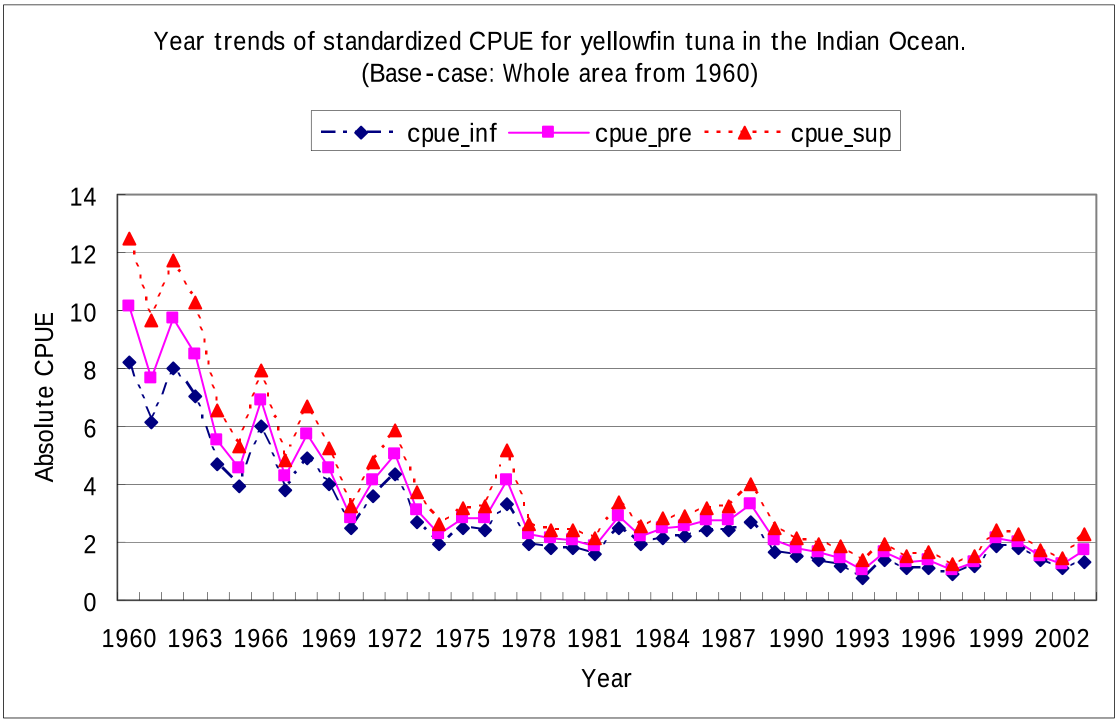


Figure 2 Trends of standardized CPUE and nominal one those are not weighted by area size in the Base-case.



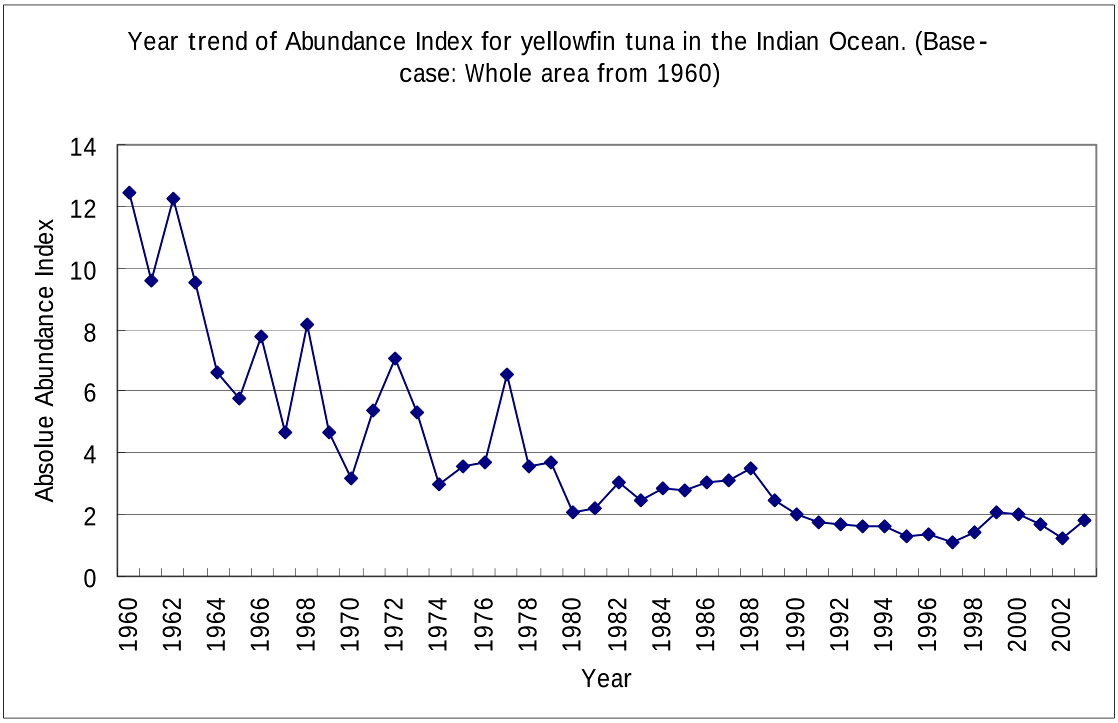
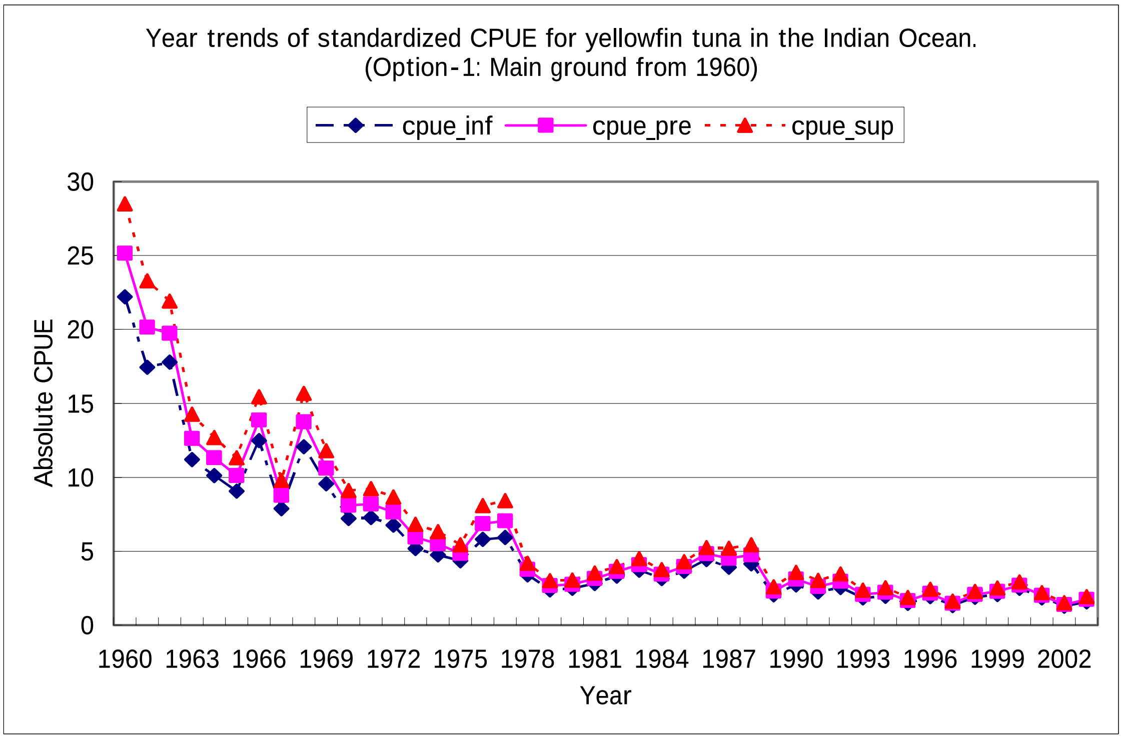


Figure 3 Absolute year trends of standardized CPUE (Dotted lines show the 95% confidence intervals) and Abundance Index (i.e. CPUE weighted by relative area-size) for yellowfin tuna in the Indian Ocean of our Base-case.



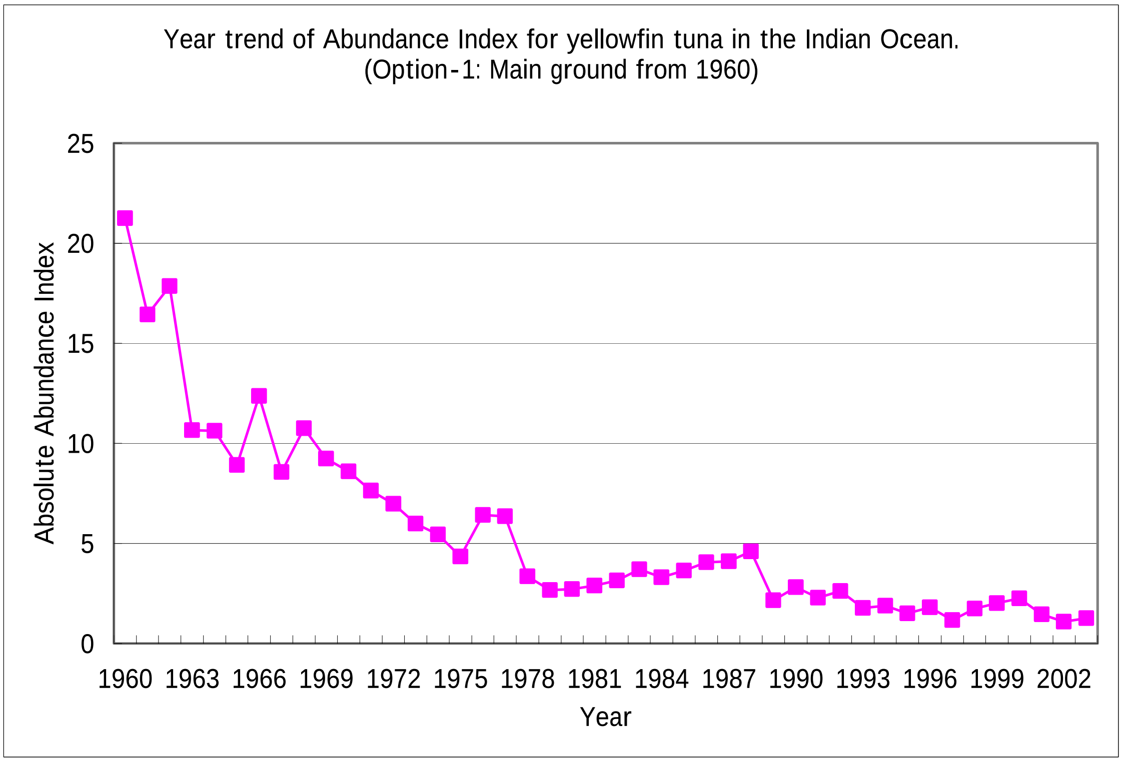
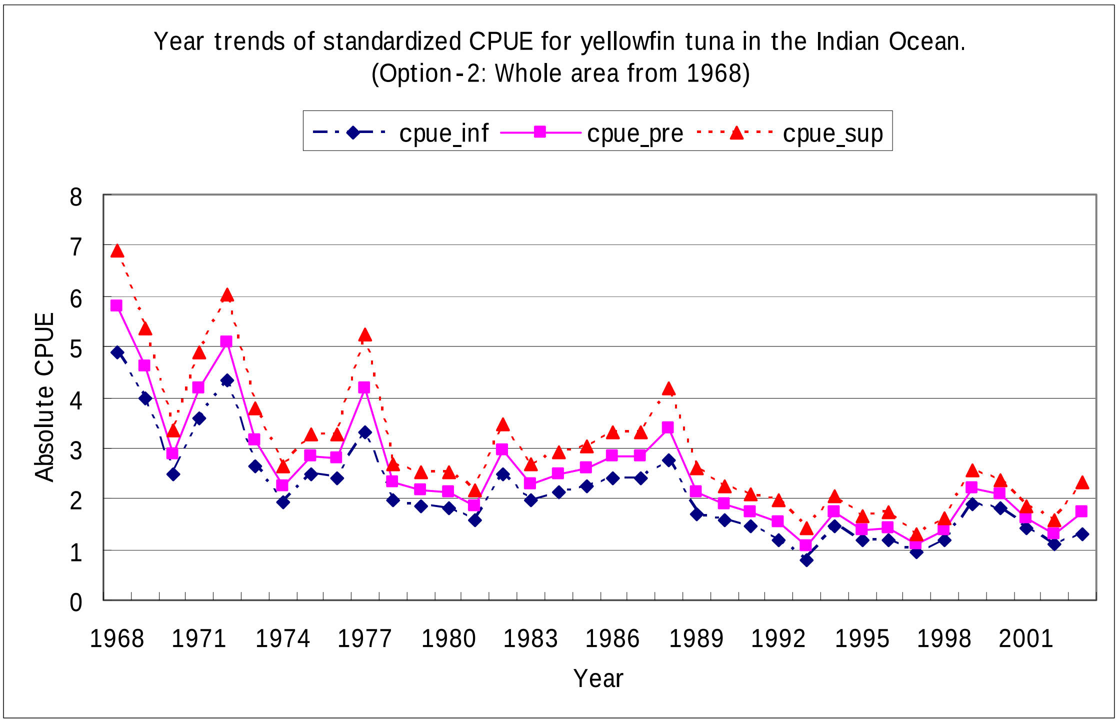


Figure 4 Absolute year trends of standardized CPUE (Dotted lines show the 95% confidence intervals) and Abundance Index (i.e. CPUE weighted by relative area-size) for yellowfin tuna in the Indian Ocean of our Option-1.



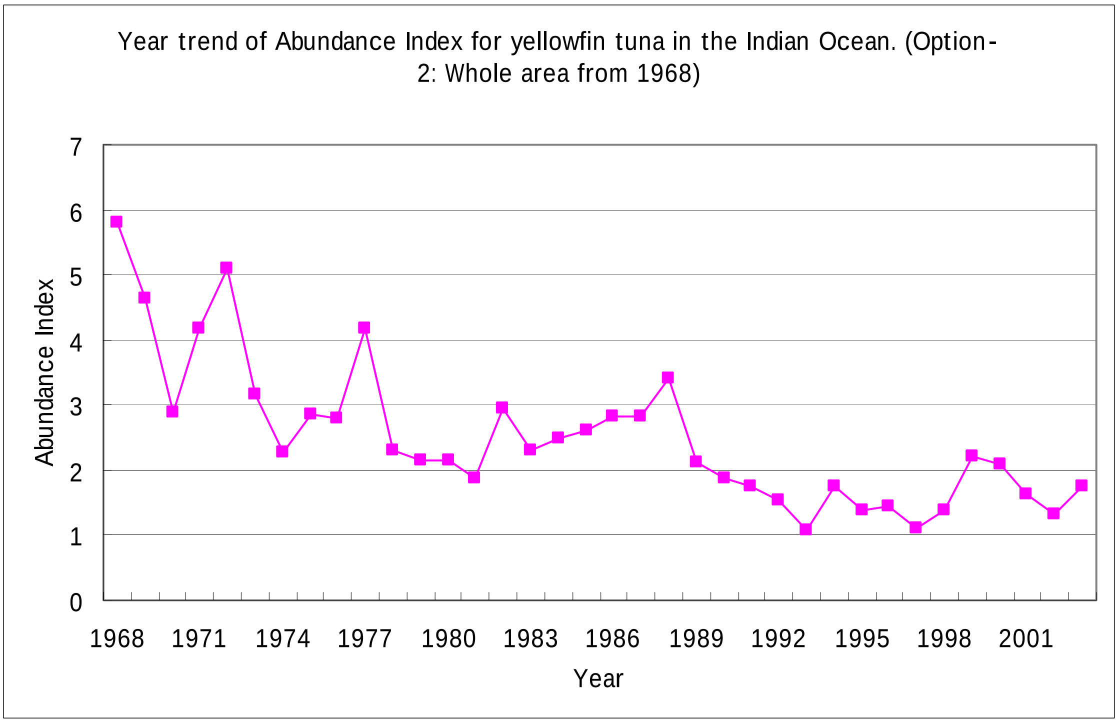


Figure 5 Absolute year trends of standardized CPUE (Dotted lines show the 95% confidence intervals) and Abundance Index (i.e. CPUE weighted by relative area-size) for yellowfin tuna in the Indian Ocean of our Option-2.

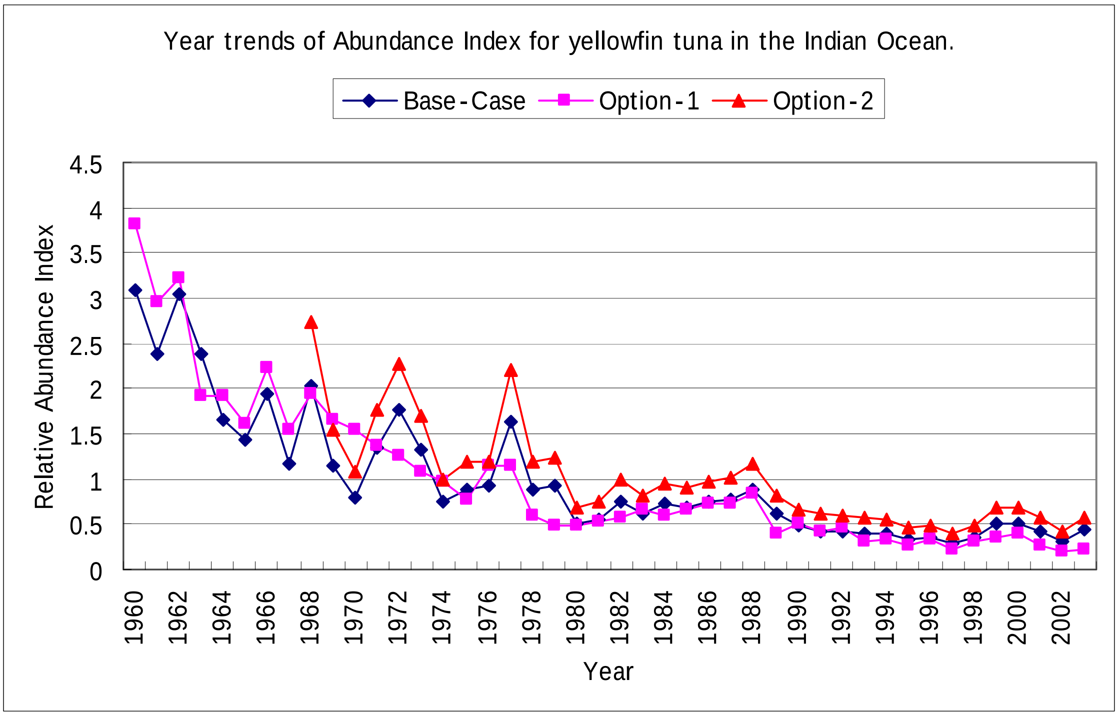


Figure 6 Relative year trends of Abundance Index (i.e. CPUE weighted by relative area-size) for yellowfin tuna in the Indian Ocean (Average value is set to be 1.0)

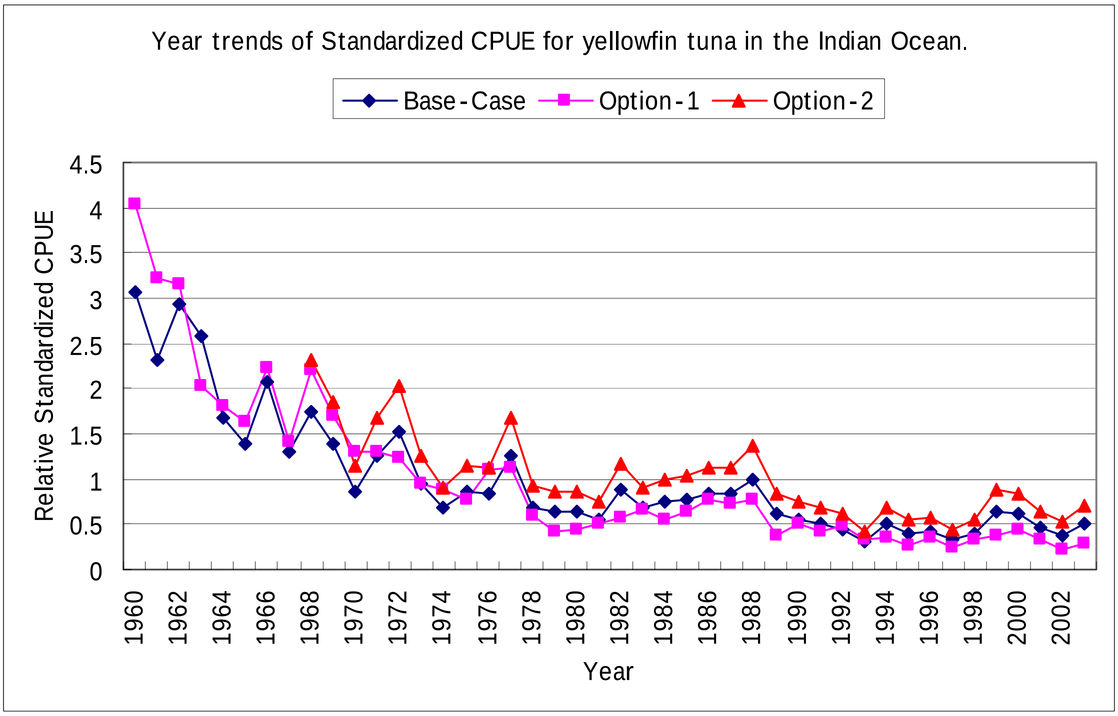


Figure 7 Relative year trends of Standardized CPUE for yellowfin tuna in the Indian Ocean (Average value is set to be 1.0.)

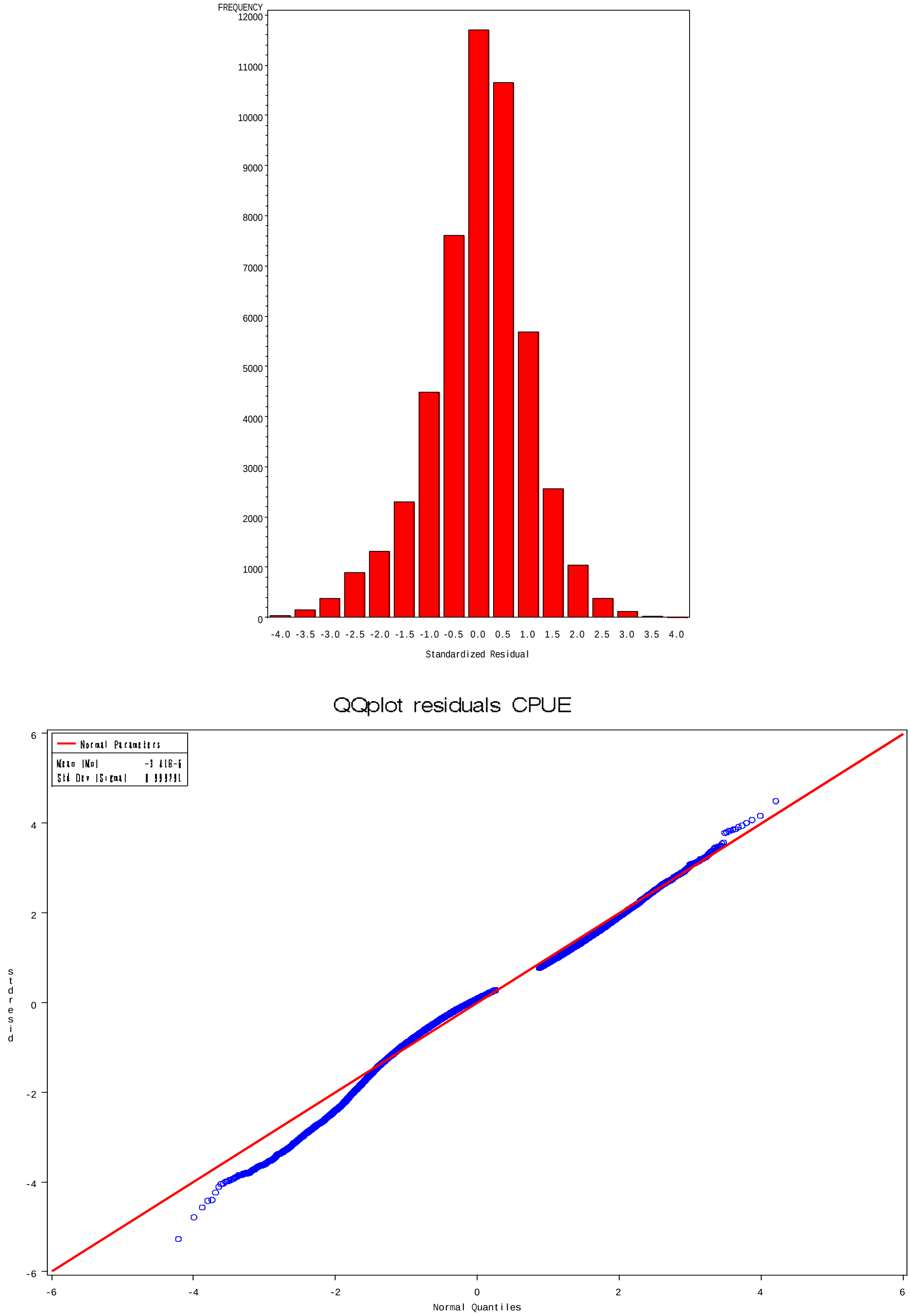
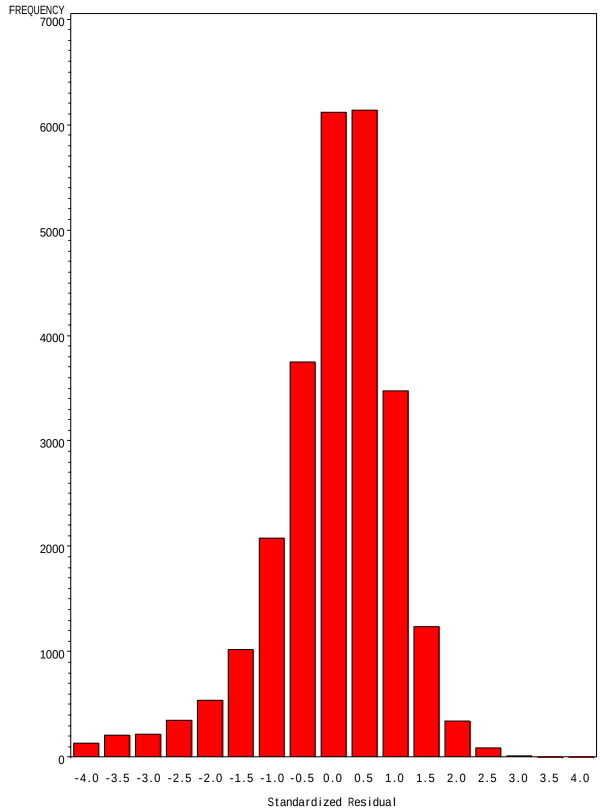


Figure 8　Histogram and QQ-plot to standardized residuals in the Base-case



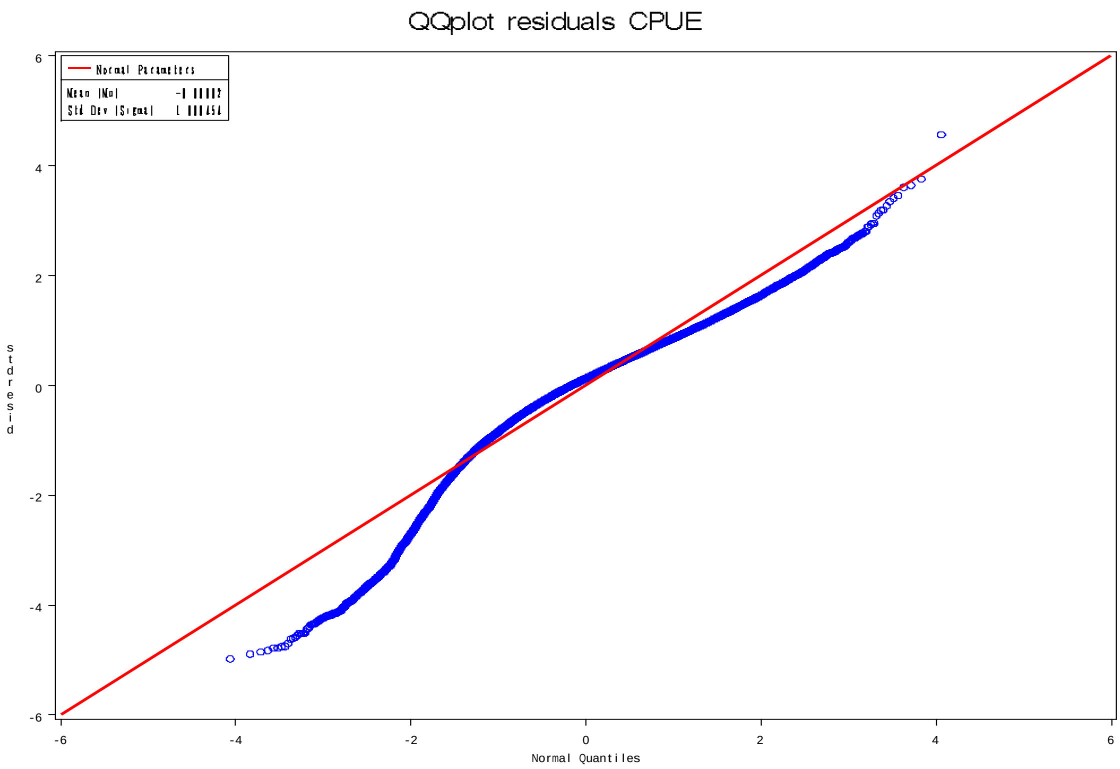
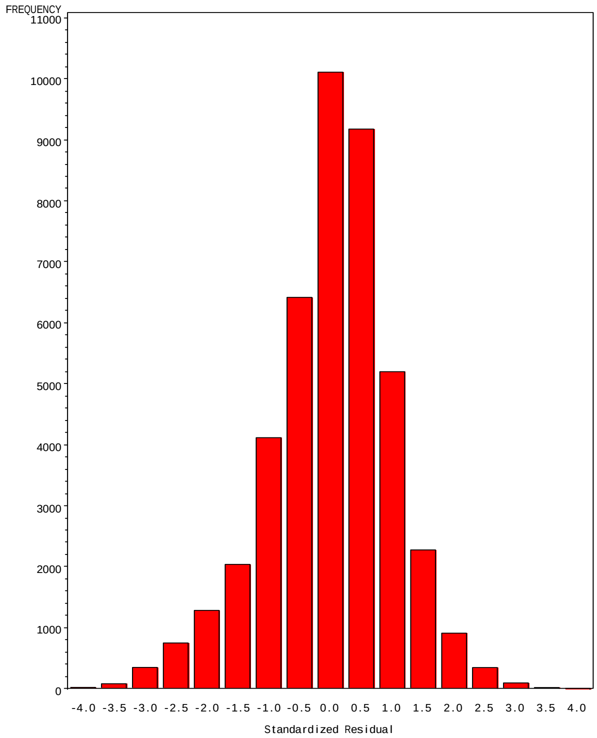


Figure 9 Histogram and QQ-plot to standardized residuals in the Option-1.



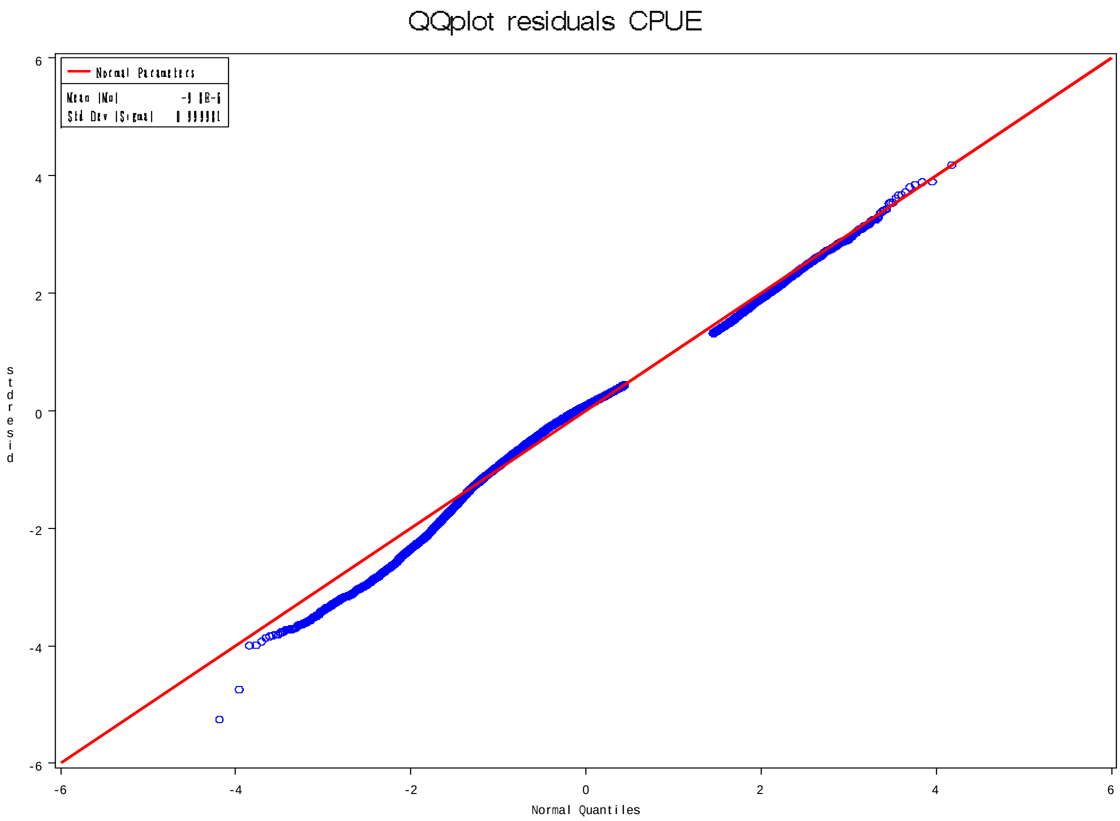


Figure 10 Histogram and QQ-plot to standardized residuals in the Option-2