CPUE Standardizations for Yellowfin Tuna Caught by Taiwanese Longline Fishery in the Indian Ocean Using Generalized Liner Model and Generalized Linear Mixed Model

Yu-Min Yeh[1] and Shu-Ting Chang[2]

[1]Graduate Institute of Environmental Management, Nanhua University, 32, Chung Keng Li, Dalin, Chiayi 62248, Taiwan

[2]Overseas Fisheries Development Council, 19, Sec. 3, Roosevelt Rd. Taipei 100, Taiwan

1 Introduction

1.1 Historical development of Taiwanese longline fishery in the Indian Ocean

Taiwan began to develop distant water tuna longline fisheries in the mid-60s. Early distant water operations targeted albacore and yellowfin for export to foreign canneries. Until the early 80s, Taiwanese tuna longline fishery expanded the ultra-low freezing technology (ULT) tuna operations. Bigeye and yellowfin are the major species caught by the ULT tuna longliners, while albacore is still a major target species for a large Taiwan fleet in the Indian Ocean longline (Haward and Bergin 2000).

YFT is among the most primary target species for longline fishing in the open seas operating in the perimeter around Indian Ocean. There was an observable change when Taiwanese longline fishing activities shifted target species from albacore to bigeye. Looking into the history of YFT longline fishing in the Indian Ocean, prior to the late 80s, the average catch recorded at lower than 10,000 mt. However, as a result of a shift of target species from albacore to bigeye, the YFT catch started increasing between 20,000 to 30,000(Chang et al. 2008) mt. It spiked at an excessive high of 80,000 mt in 1993. However, this number was not maintained; until another significant spike, which was recorded at 60,000 mt 2005. It is noteworthy that these catches were recorded as coming from fishing activities in the fishing grounds off Pakistan and Oman (Chang et al. 2008; Haward and Bergin 2000).

1.2 The yellowfin status in the Indian Ocean

The current status is total stock and spawning stock biomass above or just below MSY level (250,000-360,000 t), that is risk of overfished state. Since Fishing mortality is mostly above MSY level (1.22 – 1.75), this means overfishing is occurring. At this level of fishing pressure, it is obviously that the stock could be overfished in 3-5 years (medium term). However, observed trends between catch and Longline CPUEs are not consistent with any known theory of fishing. Probably, some major unknown factors are influencing the abundance index Targeting behavior and increased efficiency of the fleets are not totally accounted for (IOTC 2008).

1.3 Summary of the previous CPUE standardizations for Yellowfin Tuna Caught by Taiwanese Longline Fishery in the Indian Ocean

"For stock assessment purposes, the standardizations of CPUE for YFT caught by Taiwanese longline fishery in the Indian Ocean were conducted by generalized linear model (GLM) and generalized linear mixed model (GLMM), on the set by set logbooks data. The logbook data are available only from 1979; and to avoid the effects of complicated regime shifts of high catches in Oman waters and to be comparable to Japanese standardization CPUE series, the standardizations in the Taiwanese series were conducted only within the years of 1979-2008 and within tropical areas"(Chang et al. 2008). In the previous study, the rule of data extraction is to exclude the high catch composition with BET>75% and catch logged as zero or the information is entirely unavailable for YFT or ALB, due to the data coming from specific BET-targeting fishing activities; and to exclude yellowfin catch recorded at zero due to incomplete information in the data provided(Chang et al. 2008).

1.4 Purpose of the study

To provide an update of indices of abundance for yellowfin tuna from the Taiwanese longline fishery presented for the period 1979-2008. Another objective is to investigate the alternative target proxy based on the observer data which is available from 2002 to 2008.

2 Material and Method

In this study, for the base case, the researchers follows the procedure adopted in previous study(Chang et al. 2008) but with recent data updates and with new area definition.

2.1 Data set

In this study, daily set-by-set catch and effort data from the logbooks of Taiwanese longline fishery from 1979-2008 were provided by Overseas Fisheries Development Council (OFDC). In addition to this, the data on the number of hooks between floats (NHBF) were available since 1995 and the percentage of data with NHBF was about 80% of the total data from 1995 to 2008. To avoid the effects of complicated regime shifts of high catches in Oman waters and to be in comparable with Japanese standardization CPUE series, the standardizations on Taiwanese series were conducted only for years of 1979-2008 (2008 is still preliminary) and for new defined tropical areas.

2.2 Statistical models

Statistical models of GLM and GLMM were used to model the logarithm of the nominal CPUE (defined as the number of fish per 1,000 hooks) in this study. The main factors considered in this study are year, season (Jan.-Mar., Apr.-Jun., Jul.-Sep., and Oct.-Dec.), area (Areas 2 and 5), and target. The interactions between the main factors are also included in the model. The information of NHBF or hooks per basket (HPB) is usually used as target proxy in the CPUE standardization models. However, this information was only available from 1995 onwards in the logbooks of Taiwanese longline fishery. Alternative indicators were therefore adopted in the study:

1. Four categories of Bigeye catch composition defined based on the information of NHBF from observer data (1: <15%; 2:15%-25%; 3: 25%-70% 4: >70%)

2. Two categories of Albacore catch composition defined based on the information of NHBF from observer data (1: <70%; 2:>=70%).

Therefore, in this study, the researchers carried out two cases as follows:

Case	Target Proxy	Factors	Interactions
Sensitivity Case	RALB RBET	Year Season Area RALB RBET	Year*Area Season*Area Area*RALB Area*RBET
Base Case	**RYFT**	**Year Season Area RYFT**	Year*Season Year*Area Area*Season Area*RYFT

(1) GLM model: The CPUE is predicted as a linear combination of the explanatory variables.

At first, the following form was assumed as a full model.

$$\log(\mathrm{CPUE} + c) = \mu + Y + S + A + T + \text{interactions} + \varepsilon$$

where *CPUE* is the nominal CPUE of yellowfin tuna,

c is the constant value (i.e. 0.1),

μ is the intercept,

Y is the effect of year,

S is the effect of season,

A is the effect of fishing area,

T is the Target proxy,

Interactions is the interactions between main effects,

ε is the error term, $\varepsilon \sim N(0, \sigma^2)$.

Fishing areas used in this study were redefined by five new areas based on the IOTC statistics areas for yellowfin tuna in the Indian Ocean (Fig. 2):

1. Area 1: Arabian Sea;

2. Area 2: Western Indian Ocean;
3. Area 3: Mozambique Channel;
4. Area 4: Southern Indian Ocean and Atlantic-Indian Region;
5. Area 5: Bay of Bengal, Eastern Indian, and Java Sea;

(2) GLMM model: This model assumes a delta lognormal error distribution for the positive catch rates. The model fits separately the proportion of positive sets, assuming a binomial error distribution, and the mean catch rate of positive sets (at least one fish was caught) assuming a lognormal error distribution.
Estimated proportion of successful sets is assumed to be the result of r positive sets of a total n number of sets, and each one is an independent Bernoulli-type realization. The estimated proportion (ρ) is a linear function of fixed effects and interactions, by using logit function as a link between linear factor components and binomial errors. The systematic component is defined as:

$$\log\left(\frac{\rho_{Y,S,A,T}}{1-\rho_{Y,S,A,T}}\right) = \alpha_0+\alpha_Y+\alpha_S + \alpha_A+\alpha_T + \alpha_{Interactions} + \omega$$

with a binomial density:

$$\omega_{Y,S,A,T} \sim Bin(n_{Y,S,A,T}, \rho_{Y,S,A,T}\alpha_0)$$

For positive observations, which were defined as at least one yellowfin caught, the estimated CPUE rate was assumed to follow a lognormal error distribution (logCPUE) of a linear function of fixed factor and random effect interactions. The systematic component is defined as:

$$\log\left(\mu_{Y,S,A,T}\right) = \beta_0+\beta_Y+\beta_S + \beta_A+\beta_T + b_{Interactions} + z$$

with a log-normal density:

$$z_{Y,S,A,T} \sim LogNorm(\mu_{Y,S,A,T}, \sigma^2{}_{Y,S,A,T})$$

2.3 Statistical runs

This study has conducted a set of standardization runs using logbook data, by both GLMM and GLM model (Table 1). All runs only keep significant factors ($p<0.01$) in the analysis of CPUE by the effective effort. The calculation was done using GLM, GLIMMIX and MIXED procedure of SAS (Ver.9. 02). The standardized CPUE were then computed from the least square means (LSMeans) of the estimates of the year effects.

3 Results and Discussion

Since Observer data information was available only since 2002, hence alternatives were adopted as a target proxy for the whole series The researchers performed the give GLM runs on the data with the same model structure except the indicator of target factor: RYFT, RALB AND RBET. The result shows there is no obvious difference in standardized CPUE trend (Fig. 9) between base case and sensitivity case.

Table 1 and Table2 show the ANOVA tables for base case and sensitivity case separately. The model of all runs explained less than 50% of the variance, both R squares were about 40% of the variance. Distributions of the standardized residuals for the two runs are shown in Fig. 4 and Fig. 7 The distribution of the standardized residuals for both cases appear to deviate slightly from normal distribution assumption. The normal probability plots are showed in Fig. 5 and Fig. 8. The qqplots show some extent of divergence for left tail.

Relative standardized CPUEs obtained from two runs are shown in Fig. 9 and the series look very similar series. In this study, the researchers did very preliminary analysis of observer data. However, this may be worth further investigations on the observer data in the future to consider the technological effect issue in CPUE standardizations.

Table 1. ANOVA table of the selected model for base case.

ANOVA

Source	DF	Sum of Squares	Mean Square	F-value	P-value
Model	158	211443.35	1338.25	2136.90	<.0001
Error	377352	236319.35	0.63		
Corrected Total	377510	447762.70			
	R-Square	Coeff Var	Root MSE	CPUE Mean	
	0.47	103.39	0.79	0.74	
Source	DF	Type III SS	Mean Square	F-value	P-value
year	29	3318.87	114.44	182.74	<.0001
Area	1	257.00	257.00	410.37	<.0001
Season	3	356.66	118.89	189.84	<.0001
Ryft	3	149372.33	49790.78	79505.30	<.0001
year*Season	87	5500.28	63.22	100.95	<.0001
year*Area	29	1276.77	44.03	70.30	<.0001
Area*Season	3	340.73	113.58	181.36	<.0001
Area*Ryft	3	426.40	142.13	226.96	<.0001

Table 2. ANOVA table of the selected model for sensitive run.

ANOVA

Source	DF	Sum of Squares	Mean Square	F-value	P-value
Model	156	185169.03	1186.98	1705.72	<.0001
Error	377354	262593.67	0.70		
Corrected Total	377510	447762.70			
	R-Square	Coeff Var	Root MSE	CPUE Mean	
	0.41	112.15	0.83	0.74	
Source	DF	Type III SS	Mean Square	F-value	P-value
year	29	3539.24	122.04	175.38	<.0001
Area	1	381.08	381.08	547.62	<.0001
Season	3	358.79	119.60	171.87	<.0001
Ralb	1	4547.38	4547.38	6534.70	<.0001
Rbet	3	129095.07	43031.69	61837.70	<.0001
year*Area	29	1641.46	56.60	81.34	<.0001
year*Season	87	5547.30	63.76	91.63	<.0001
Area*Rbet	3	208.05	69.35	99.66	<.0001

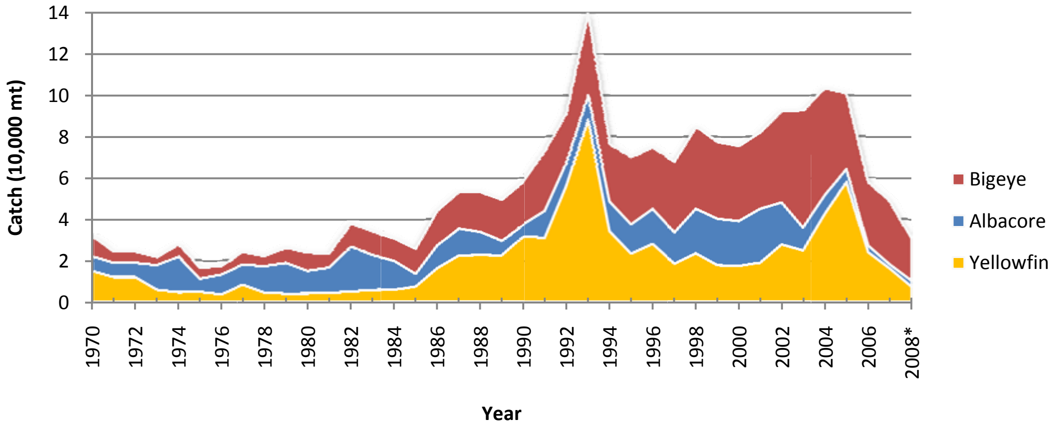

Figure 1. Nominal catches (mt) of main target species caught by Taiwanese longline fishery in the Indian Ocean over the period 1970 to 2008.

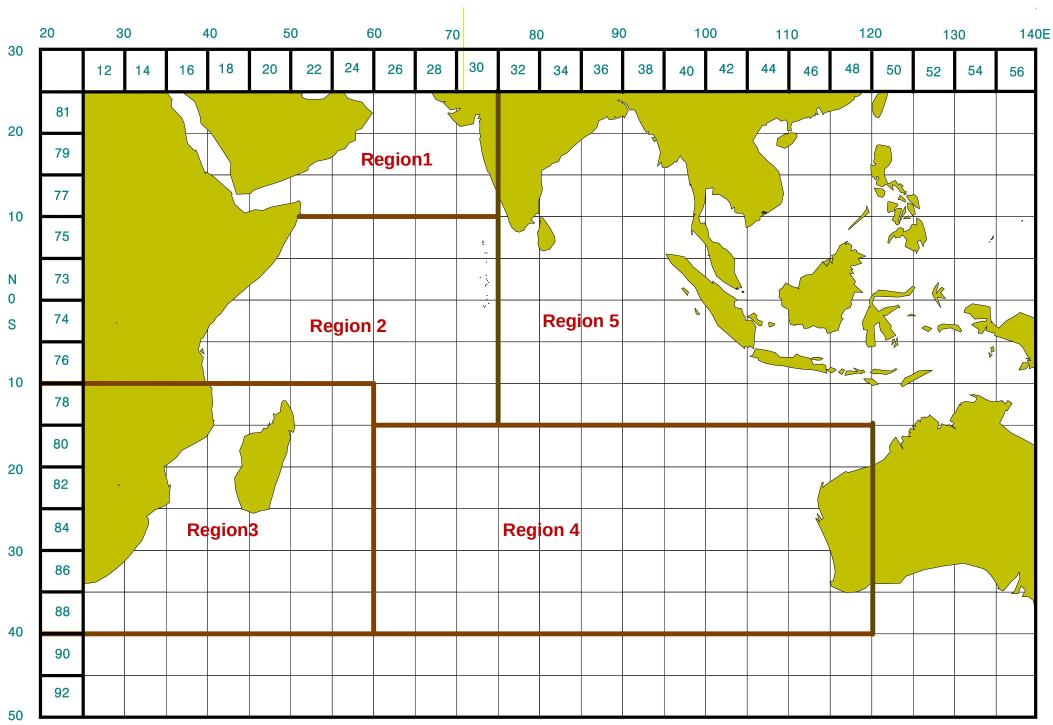

Figure 2. New area stratification used for the standardization of CPUE for yellowfin tuna in the Indian Ocean in 2009.

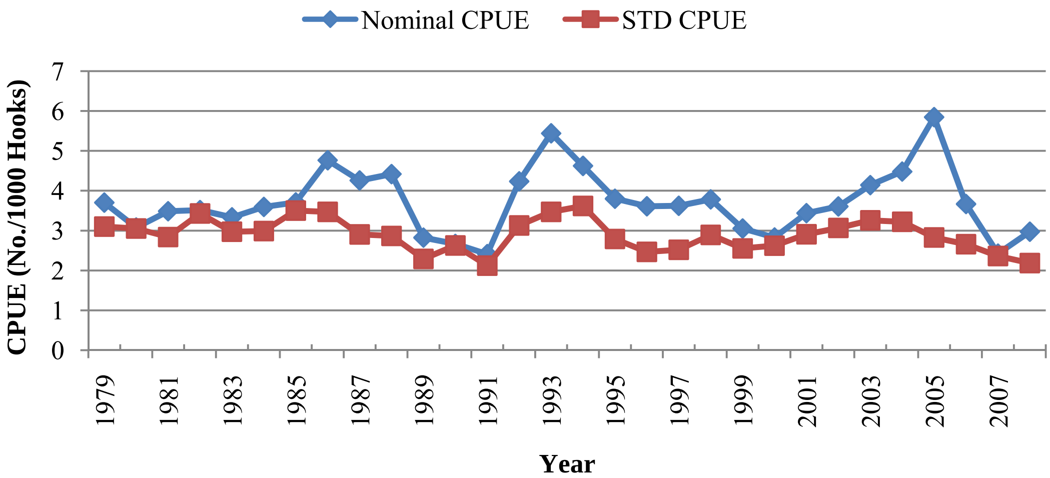

Figure 3. Nominal and Standardized CPUE series for base case.

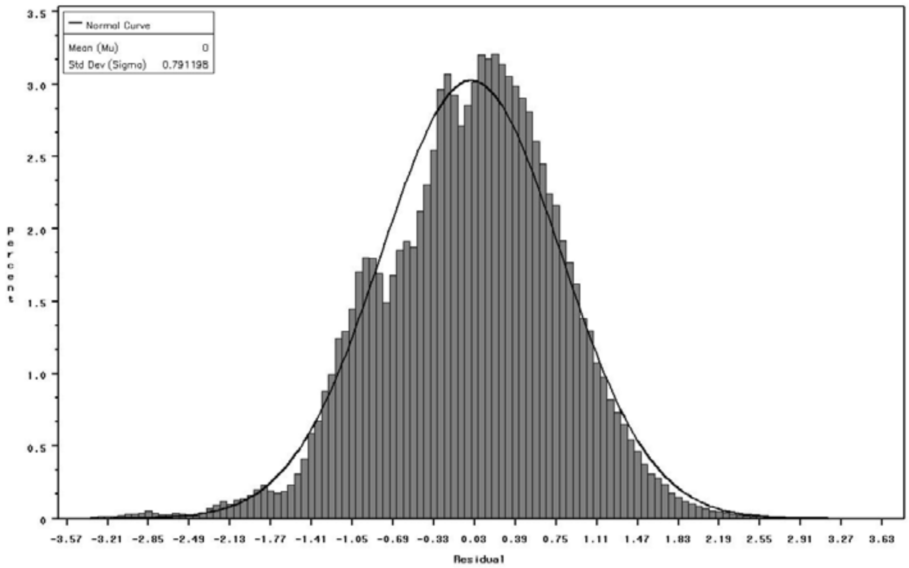

Figure 4. The residuals distribution for base case.

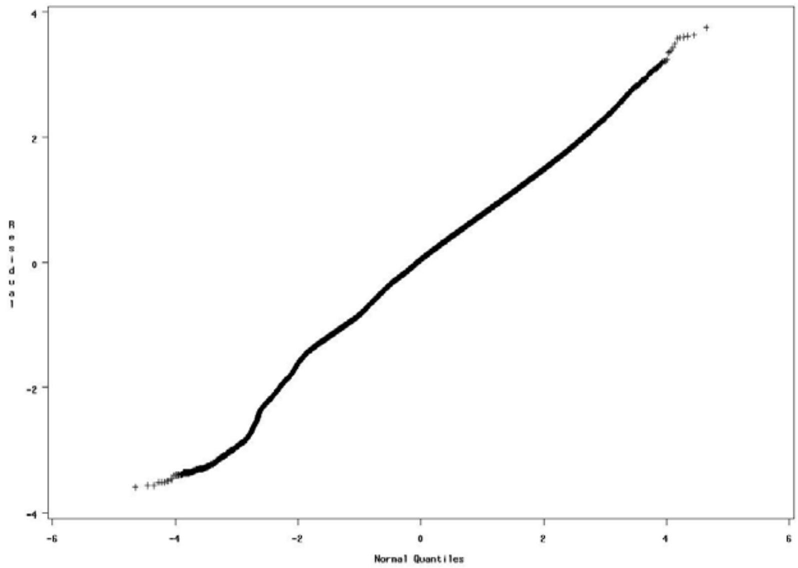

Figure 5. The QQPlot for base case.

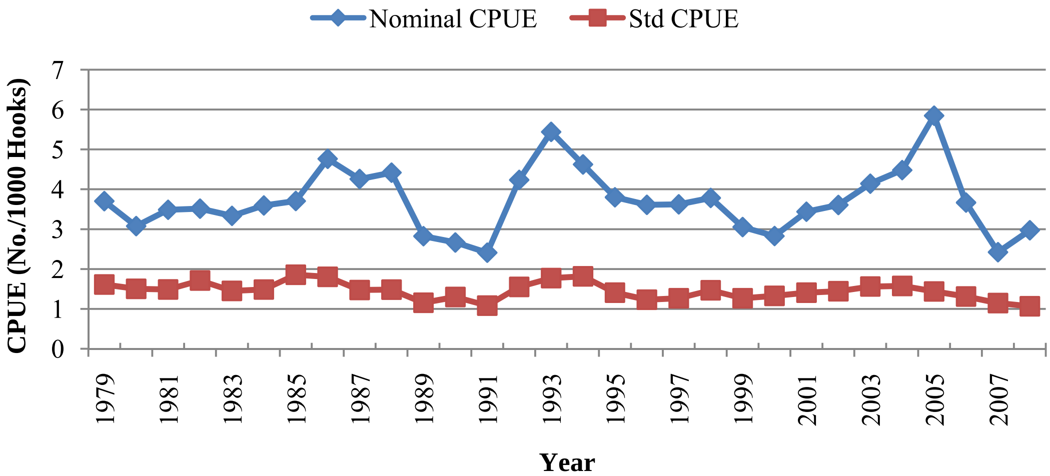

Figure. 6. Nominal and Standardized CPUE series for sensitivity case.

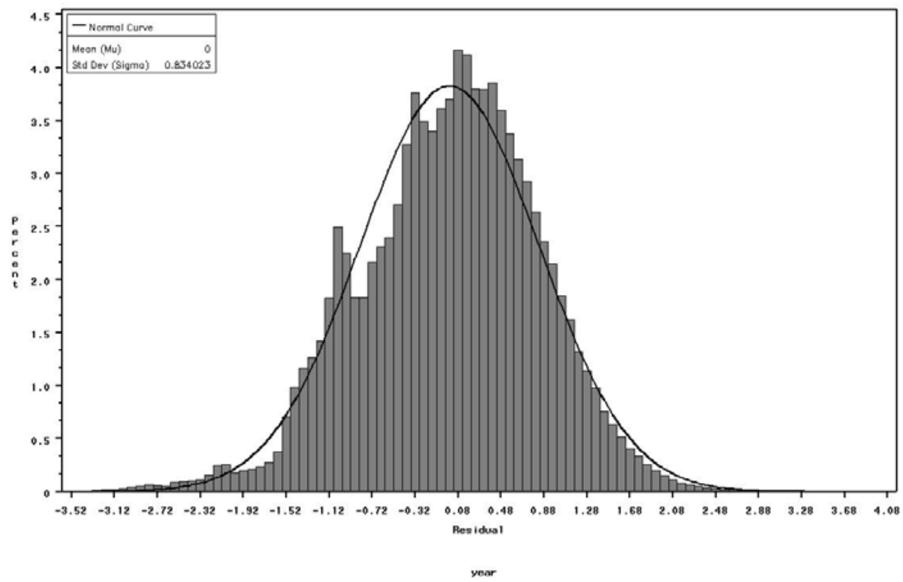

Figure. 7 The residuals distribution for sensitivity case.

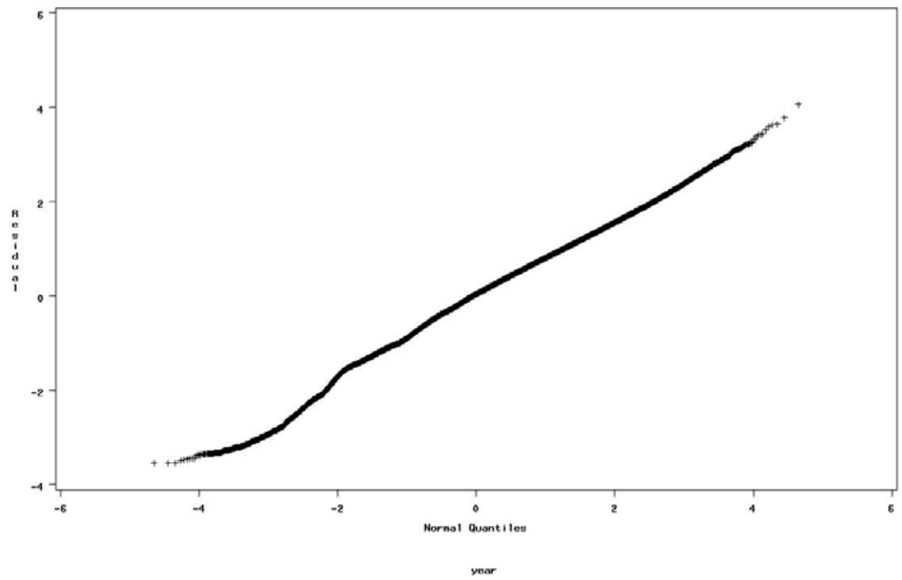

Figure. 8 The QQPlot for sensitivity case.

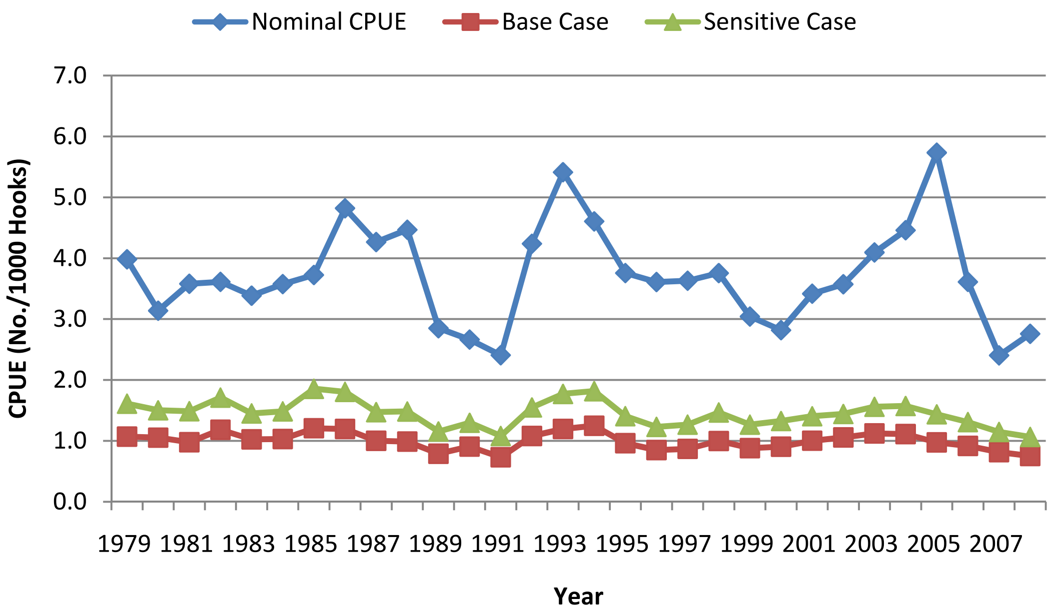

Figure 9. Nominal and Standardized CPUE series for two runs for comparison.

4 References

Chang S-K, Liu H-I, Chang S-T (2008) CPUE standardizations for yellowfin yuna caught by taiwanese deep sea longline fishery in the tropical Indian Ocean using generalized linear model and generalized linear mixed model. Paper presented at the The indian tuna commission working party on tropical tunas in 2008, Bangkok, Thailand.

Devaraj M (1982) A critique on Indian Ocean fisheries development. 8(2):97-123.

Haward M, Bergin A (2000) Taiwan's distant water tuna fisheries. 24(1):33-43.

IOTC (2008) Report of the Tenth Session of the IOTC Working Party on Tropical Tunas, in The Tenth Session of the IOTC Working Party on Tropical Tunas, Bangkok, Thailand.