# STANDARDIZED CPUE FOR YELLOWFIN TUNA (*THUNNUS ALBACARES*) OF THE JAPANESE LONGLINE FISHERY IN THE INDIAN OCEAN BY GENERALIZED LINEAR MODELS (GLM) (1960-2000)

Hiroshi SHONO, Hiroaki OKAMOTO and Tom NISHIDA

*National Research Institute of Far Seas Fisheries*

*7-1, 5-chome Orido Shimizu-shi, 424-8633, Japan*

**ABSTRACT**

*We performed the CPUE standardization for yellowfin tuna in the Indian Ocean from 1960 to 2000 using the updated Japanese longline catch and effort data which were aggregated by month, 5x5 degree square and the number of hooks between floats (NHF) by the typical generalized linear models (GLM) called CPUE model with log-normal error. The main purpose of this analysis is to make the CPUE index as an input for age specific production models (ASPM). We incorporated the six explanatory variables, YEAR, MONTH, AREA, NHF, SST and SOI, as the main effect and/or two-way interactions into the GLM analyses. We computed the standardized yellowfin CPUE and abundance-index (i.e. area-weighted CPUE) using the statistical F-test based on the value of deviance. These two indices selected show very similar trends, decrease drastically in 1960s and 1970s, and after that seem to be rather stable.*

## INTRODUCTION

There are several studies regarding CPUE standardization of tropical tuna such as bigeye and yellowfin in the Indian Ocean using generalized linear models (GLM) (Nishida, 2000: Okamoto *et al.*, 2001). However, full-scaled stock assessment about yellowfin tuna has not been done since 1995 because of the lack of the size data. Therefore, we carried out the CPUE standardization of yellowfin tuna in the Indian Ocean using updated Japanese longline data by the GLM and estimated CPUE index (i.e. year trend of abundance) as an input for an age-specific production model (ASPM) (Nishida and Shono, 2002).

## 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 from 1952 to 1974 were used, although it did not include the information on NHF. Then, we assume that the NHF is 6.

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

log(CPUE+constant) = INTERCEPT + YEAR + MONTH + AREA + GEAR

+ SST + SOI + INTERACTION + ERROR, ERROR~ N $(0, s^2)$ (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,

GEAR: effect of NHF,

SST: effect of sea surface temperature,

SOI: effect of southern oscillation index,

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 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 -

$$\begin{aligned}\log(CPUE_{ijkl}+0.1) &= INTRECEPT + YEAR_i + MONTH_j + AREA_k + GEAR_l + SST \\ &+ SOI + (YEAR*AREA)_{ik} + (MONTH*AREA)_{jk} + (MONTH*GEAR)_{jl} \\ &+ (AREA*GEAR)_{kl} + (AREA*SST)_k + (AREA*SOI)_k + ERROR_{ijkl}\end{aligned} \quad (2)$$

where

i (YEAR): 1960-2000,

j (MONTH): 1-12,

k (AREA): 1-6,

l (GEAR): 1-4 (class 1: 5-6, class 2: 7-10, class 3: 11-14, class 4: 15-24).

Remark) SST and SOI were incorporated as a continuous variable.

YEAR, MONTH, AREA, GEAR, SST and SOI were incorporated as the main effect. We used the data from 1960 to 2000. The area stratification is shown in Figure 1. Considering the distributions of effort (the number of hooks per 1000) and CPUE of yellowfin tuna in the Indian Ocean (shown in Figure 2 and 3), main fishing ground was divided into six areas for the analysis. The number of branch lines between floats (NHF) was divided into 4 classes, though the NHF was fixed to 6 before 1974 because of no information about NHF, and incorporated into the model as an effect of gear.

We also added the effect of SST and SOI to the model as environmental factors that may have affected the yellowfin stock status and recruitment. Although the index of Indian oscillation index (IOI) is available from 1972 to 2000, SOI index was used instead of IOI in this analysis because of the problem about missing data and/or the possible autocorrelation problem between SOI and IOI.

**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 main effect of SOI was deleted from the full model.) Significant level was assumed to be one percentage. The results of ANOVA are shown in Table 1.

- final model -

E[log(CPUE+0.1)] = INTERCEPT + YEAR + MONTH + AREA + GEAR + SST + (YEAR*AREA) + (MONTH*AREA) + (MONTH*GEAR) + (AREA*GEAR) +(AREA*SST) + (AREA*SOI) (3)

**CPUE and Abundance Index**

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

$$CPUE_i = \exp\{INTERCEPT + YEAR_i + (\overline{YEAR*AREA})_{i\cdot}\} - 0.1,\ (i = 1960,\cdots,2000) \quad (4)$$

where

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

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.

$$CPUE_{ik} = \exp\{INTERCEPT + YEAR_i + AREA_k + (YEAR * AREA)_{ik} + (\overline{MONTH * AREA})_{\cdot k} + (\overline{AREA * GEAR})_{k\cdot} + (\overline{AREA * SST})_k + (\overline{AREA * SOI})_k\} - 0.1, \quad (i = 1960, \cdots, 2000 : k = 1, \cdots, 6) \tag{5}$$

where

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

$$(\overline{AREA * GEAR})_{k\cdot} = \frac{1}{N_l} \sum_{l=1}^{N_l} (AREA * GEAR)_{\cdot l} \qquad \text{etc.}$$

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.

$$AbundanceIndex_i = \sum_{k=1}^{N_k} w_k \cdot CPUE_{ik}, \quad (i = 1960, \cdots, 2000) \tag{6}$$

where

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

The relative area size is shown in Table 2.

**RESULTS AND DISCUSSIONS**

Figure 4 shows the comparison of the trends between standardized and nominal CPUE. 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 5 shows the standardized CPUE trend and its ninety-five percent confidence intervals. Its confidence intervals are rather narrow, especially after the period of 1980s.

Figure 6 shows the trends of standardized CPUE (that is not weighted by area size) and abundance index (that is weighted by area size). These two trends are almost same except the period of 1960s. However, in order to estimate the present/future stock status more precisely, we decided to utilize this standardized abundance index as an input for an age-specific production model (ASPM) (Nishida and Shono, 2002). The distribution of overall residual in the final model is shown in Figure 7. Judging from the distributed pattern, the log-normal model we used this time seems to be appropriate.

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. Therefore, it seems to be needed to consider the alternative statistical methods such as tree regression models (we are now trying to calculate by the tree model.) and discuss those CPUE trends from the biological point of view.


# ACKNOWLEDGEMENT

We acknowledge Mr. Takayuki MATSUMOTO, National Research Institute of Far Seas Fisheries, for giving us the useful environmental information about SOI and IOI.

# 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.

OKAMOTO, H., MIYABE, N. AND INAGAKE, D. (2001): Interpretation of high catch rates of bigeye tuna in 1977 and 1978 observed in the Japanese longline fishery in the Indian Ocean. IOTC/WPM/01/07, 25p.

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


*Table 1. Results of ANOVA for the final selected model in the GLM analysis.*

| Source | DF | Sum of Squares | Mean Square | F value | Pr > F |
|---|---|---|---|---|---|
| Model | 374 | 52639.8452 | 140.7482 | 96.66 | <.0001 |
| Error | 41129 | 59888.9991 | 1.4561 | | |
| Corrected Total | 41503 | 112528.8443 | | | |

| R-Square | Coeff Var | Root MSE | logCPUE Mean |
|---|---|---|---|
| 0.46779 | 113.6589 | 1.2067 | 1.061686 |

| Source | DF | Type II SS | Mean Square | F value | Pr > F |
|---|---|---|---|---|---|
| year | 40 | 5921.477046 | 148.036926 | 101.66 | <.0001 |
| month | 11 | 699.874007 | 63.62491 | 43.69 | <.0001 |
| area | 5 | 1780.47347 | 356.094694 | 244.55 | <.0001 |
| gear | 3 | 286.062939 | 95.354313 | 65.48 | <.0001 |
| sst | 1 | 5573.618928 | 5573.618928 | 3827.7 | <.0001 |
| year*area | 200 | 2878.887656 | 14.394438 | 9.89 | <.0001 |
| month*area | 55 | 2837.212016 | 51.585673 | 35.43 | <.0001 |
| month*gear | 33 | 311.779955 | 9.447877 | 6.49 | <.0001 |
| area*gear | 15 | 1058.99164 | 70.599443 | 48.48 | <.0001 |
| sst*area | 5 | 1639.9895 | 327.9979 | 225.25 | <.0001 |
| soi*area | 6 | 38.171414 | 6.361902 | 4.37 | 0.0002 |

| Source | DF | Type III SS | Mean Square | F value | Pr > F |
|---|---|---|---|---|---|
| year | 40 | 3711.794817 | 92.79487 | 63.73 | <.0001 |
| month | 11 | 577.668603 | 52.515328 | 36.07 | <.0001 |
| area | 5 | 2731.943081 | 546.388616 | 375.23 | <.0001 |
| gear | 3 | 106.305055 | 35.435018 | 24.34 | <.0001 |
| sst | 1 | 675.654006 | 675.654006 | 464.01 | <.0001 |
| year*area | 200 | 2878.887656 | 14.394438 | 9.89 | <.0001 |
| month*area | 55 | 2837.212016 | 51.585673 | 35.43 | <.0001 |
| month*gear | 33 | 311.779955 | 9.447877 | 6.49 | <.0001 |
| area*gear | 15 | 1058.99164 | 70.599443 | 48.48 | <.0001 |
| sst*area | 5 | 1639.9895 | 327.9979 | 225.25 | <.0001 |
| soi*area | 6 | 38.171414 | 6.361902 | 4.37 | 0.0002 |

Table 2. Area Index for weighting.

| Area | Scaled size of each area | Area index for weighting |
|---|---|---|
| 1 | 1031 | 0.074737224 |
| 2 | 2719 | 0.197100399 |
| 3 | 3206 | 0.232403045 |
| 4 | 3260 | 0.236317506 |
| 5 | 2463 | 0.17854295 |
| 6 | 1116 | 0.080898876 |
| Total | 13795 | 1 |

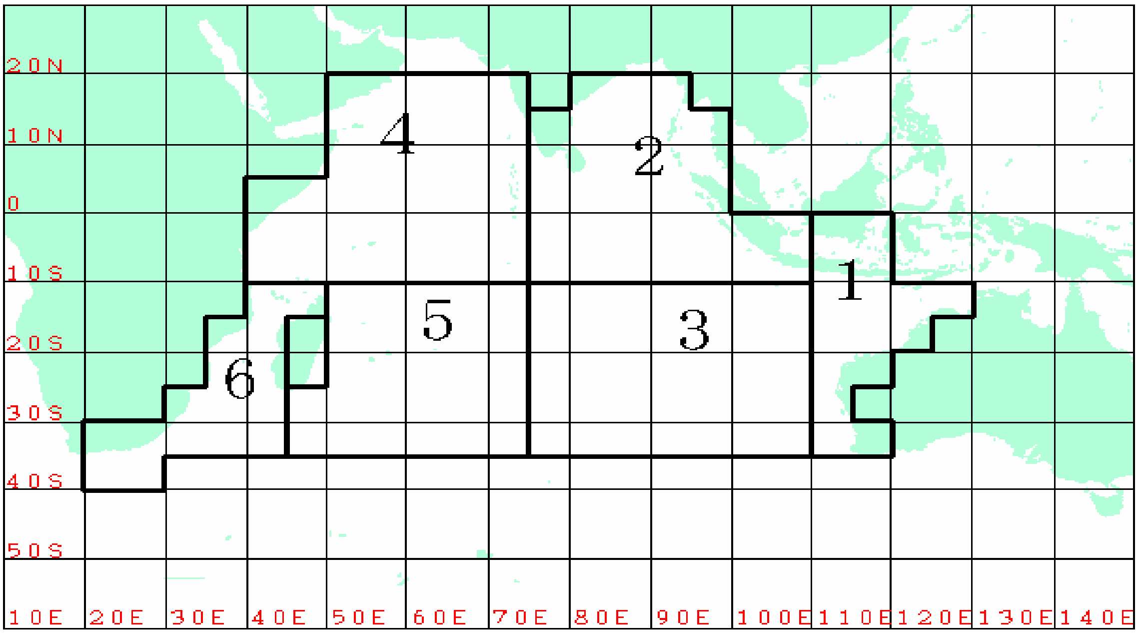


***Figure 1.*** *Area stratification used for the standardization of yellowfin CPUE.*

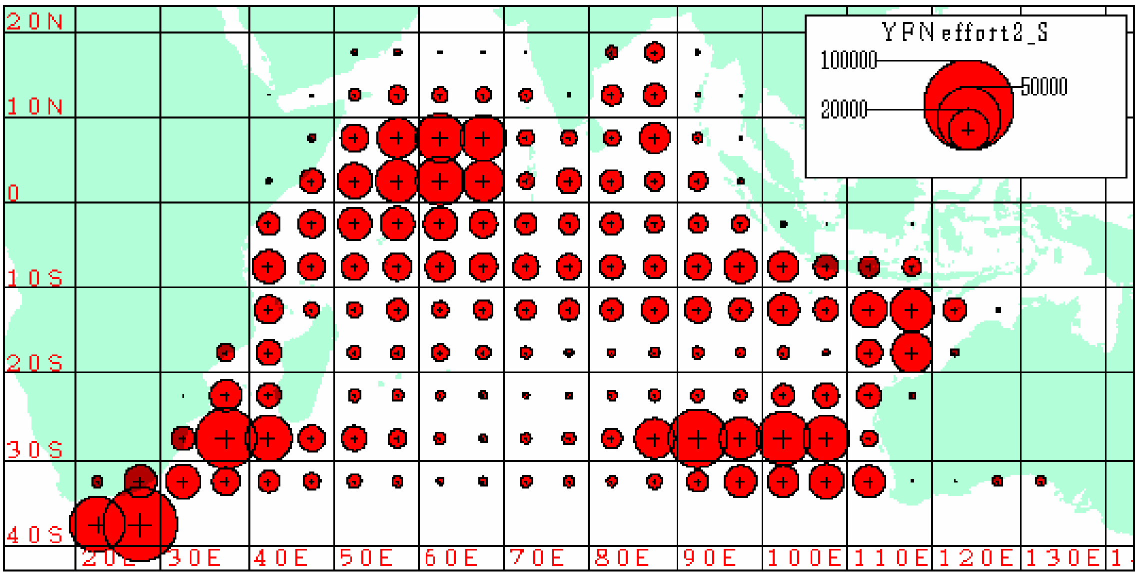


***Figure 2.*** *Geographical distribution of Japanese longline effort (in 1000 hooks) in the Indian Ocean, 1960-2000.*

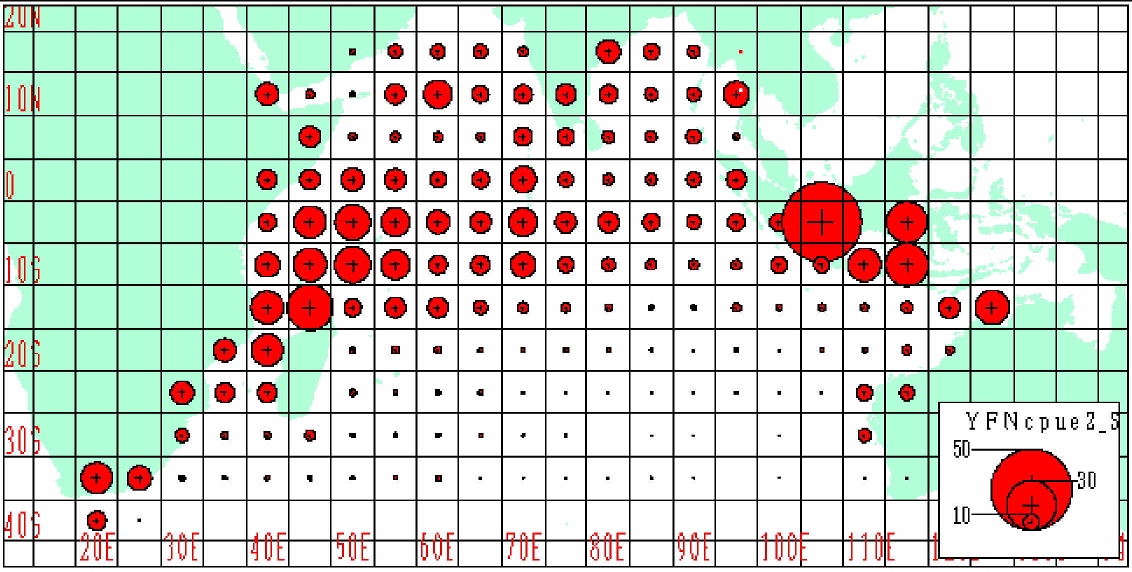


***Figure 3.*** *Nominal CPUE (catch number per 1000 hooks) for yellowfin tuna in the Indian Ocean, 1960-2000.*

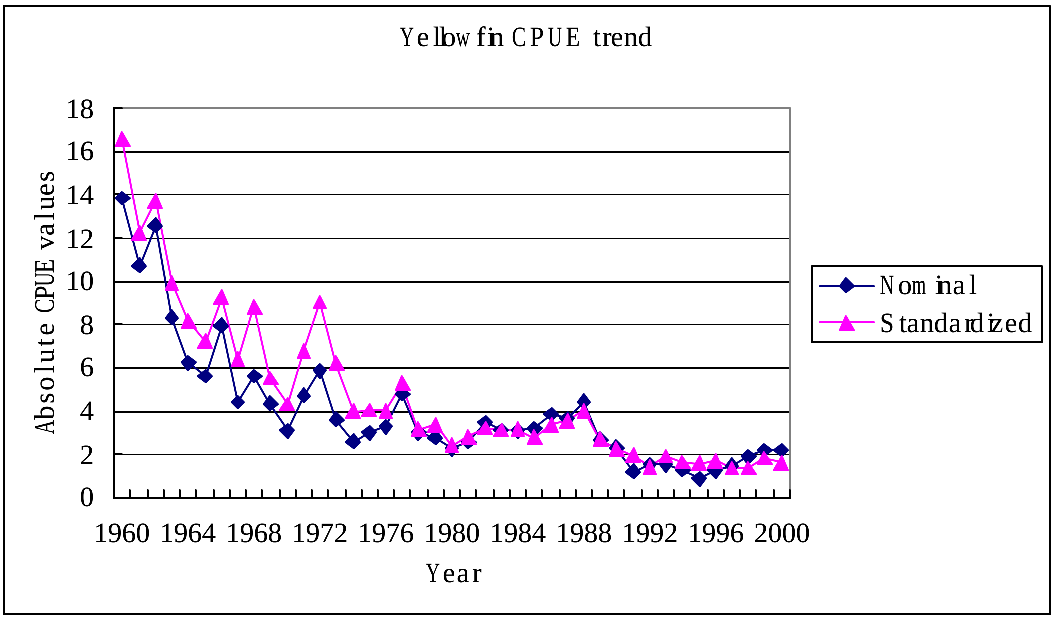


*Figure 4. Trends of standardized CPUE and nominal one those are not weighted by area size.*

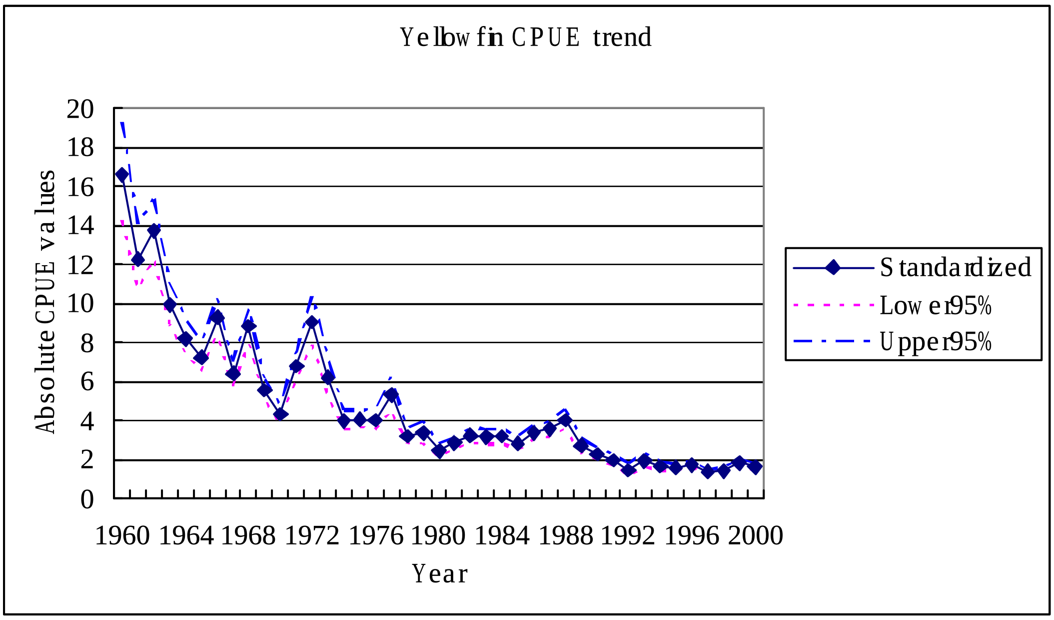


*Figure 5. Trends of standardized CPUE and its 95% confidence intervals that those are not weighted by area size.*

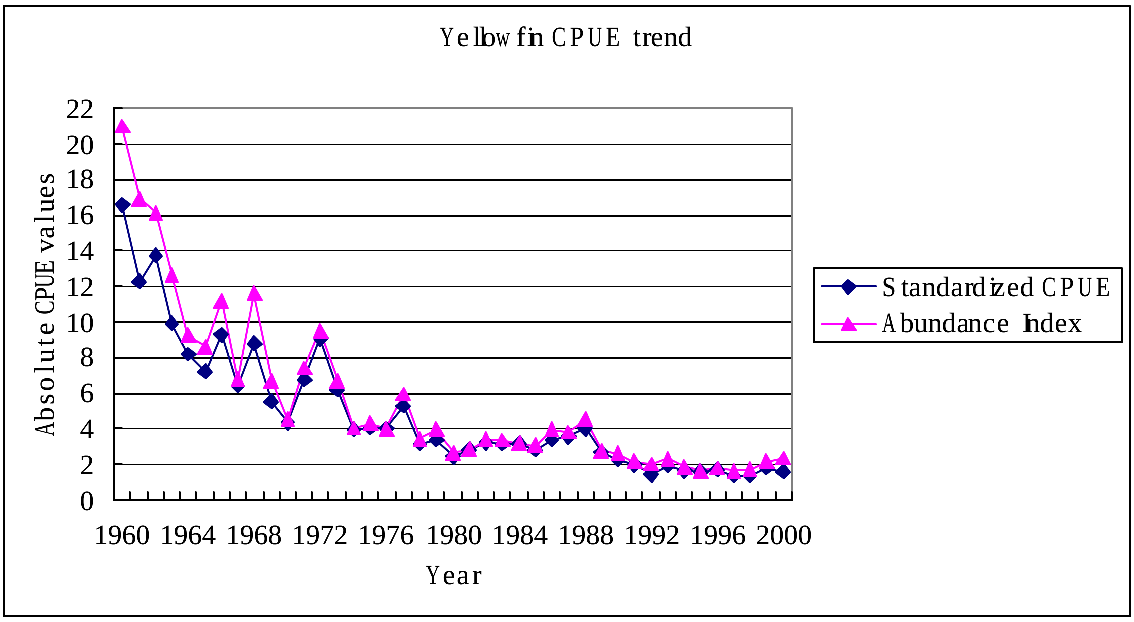


*Figure 6. Trends of standardized CPUE (that is not weighted by area size) and standardized abundance index (that is weighted by area size).*

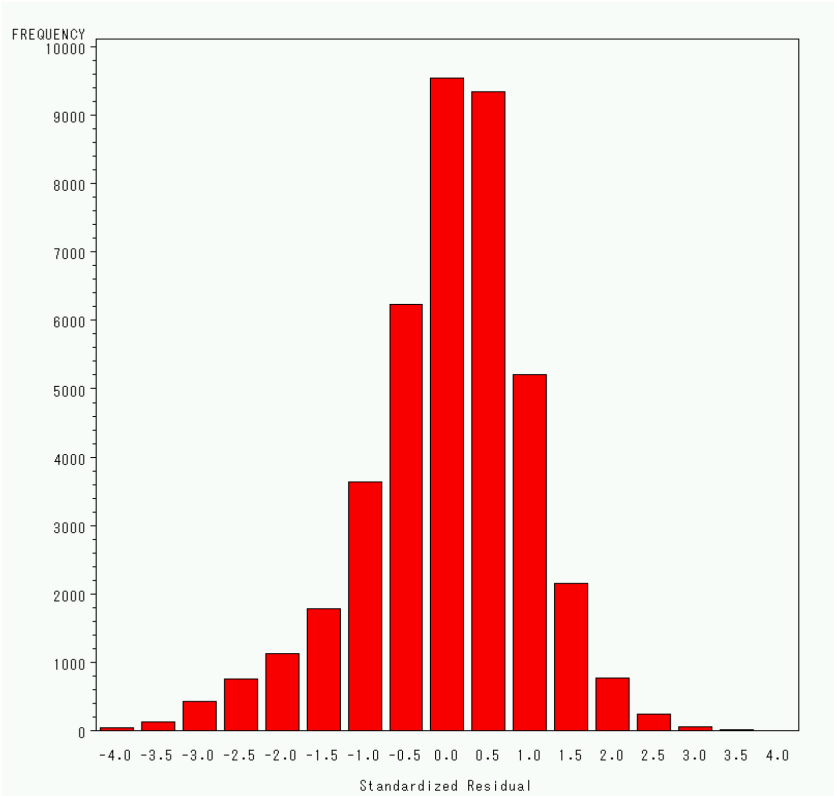


*Figure 7. Histogram to standardized residuals in this analysis, 1960-2000*