# Reviews and prospects on approaches reflecting actual dynamics of Taiwanese longline fisheries in CPUE standardizations when number of hook per basket information not available

- THE TREATMENTS OF CLASSIFICATIONS AND TARGETING FOR THE TAIWANESE LONGLINERS IN CPUE STANDARDIZATIONS -

Tom Nishida [1/] , Ying-Chou Lee [2/], Chien-Chung Hsu [3/] and Shui-Kai (Eric) Chang [4/]

*1/ National Research Institute of Far Seas Fisheries, Shizoka, Japan (tnishida@affrc.go.jp)*
*2/ Institute of Fisheries Science, College of Life Science, National Taiwan University, Taipei, Taiwan (i812@ntu.edu.tw)*
*3/ Ocean Research Institute, National Taiwan University, Taipei, Taiwan (hsucc@ntu.edu.tw)*
*4/ Fisheries Research Institute, Fisheries Agency, Taipei, Taiwan (skchang@ms1.fa.gov.tw)*

*August,, 2004*

**Contents**



---

*Submitted to the Working Party on the first TeMPerate tuna working group (WPTMP) meeting (Aug 2-5, 2004), Shizuoka, Japan, co-organized by Indian Ocean Tuna Commission (IOTC) and National Research Institute of Far Seas Fisheries (NRIFSF).*

*(*) Originally from IOTC/WPTT/2004/10*

## 1. Introduction

In the tuna longline (LL) CPUE standardization by the GLM and other related methods, the effect of the number of hook-per-basket (NHB) is one of the most important factors because the NHB (LL gear configurations) significantly affect target species and species compositions of the catch. Thus, it reflect actual changes and dynamic of the LL fisheries. Thus its effect considerably influences resultant standardized CPUE. Therefore the most ideal approach is to use several classes of the NHB information as frequently applied in the Japanese LL CPUE standardization (Okamoto *et al,* 2004 and many others).

However for the Taiwanese LL, such information are available only after 1995. Even after 1995, only around 40% of the LL data has that information. Hence, for those of the Taiwanese LL data without NHB information, several methods have been developed to conduct more accurate CPUE standardization as shown in Table 1. Merits and demerits of these methods have been reported in various meetings in the past (for example, IOTC 2002-2003). But, we don't see any global discussion to now. That is the major reason why we summarize and review available methods in this paper to be discussed in the IOTC/WPM meeting in 2004 in order to improve the existence methods and/or development new methods.

## 2. Reviews

We review the available methods. Table 1 summarizes category of the methods, authors, and titles of the papers, method types, and description of the methods, merits, demerits and actual applications. Based on this review, it was recognized that there are three major categories to reflect the actual LL fisheries in the GLM or other related methods , i.e., (A) classification (separation) of LL type, (B) Refection of the targeting and (C) others(Fuzzy and GAM). Details of these approaches of (A), (B) and (C) are described in Appendix A, B and C.

Table 1 Summary of various approaches to reflect the actual LL Taiwanese fisheries in the CPUE standardization by GLM or other related methods when number of hook per basket information are not available.

| Type | Authors | year | Titile | Method | Outline | merits | Demerits | Application |
|---|---|---|---|---|---|---|---|---|
| (A) CLASSIFICATION | (A1) Chang et al | 1993 | An Alternative Procedure to Segregate Mixed Longline Catch Data | cluster analyses | Separation of regular & deep LL by cluster analyses | Using statistical methods to decide the number and rationality of grouping and to separate into groups | The process is somewhat complicated. Not yet been tested on set by set data | Chang et al ( 1993 ) |
| | (A2) Lin | 1998 | The relationship between Taiwanese longline fishing patterns and catch compositions in the Indian Ocean | catch Ratio | ALB ratio = ALB/(ALB+BET) were mapped in ALB, BET & MIXED areas to investigate its characteristics. No separation works done. | Easy to compute. Classification accuracy might be high in the ALB Fishing grounds. | Effective only in ALB fishing grounds | (A2a) Hsu & Liu (2000) (A2b) Hsu et al (2001) (A3)Chen (1998) |
| | (A3) Chen | 1998 | Stock assessment on the Indian Ocean albacore tuna | | Separate the regular LL using the Lin's ALB ratio by month | | | |
| | (A4) Lee et al | 2004 | Separation of the Taiwanese regular and deep tuna longliners in the Indian Ocean using bigeye tuna catch ratios | | BET ratio: Regular LL for $0.8 \le BET/(BET+ALB) \le 1$ and Deep LL for $0 \le BET/(BET+ALB+SWO) \le 0.40$ | Calcification accuracy is about 70%. Effective in the ALB fishing grounds. | Ineffective if ALB catch are low. Classification accuracy is uncertain if fisheries patterns are different from those in the learning data set | Lee et al (2004) Okamoto et al (2004) |
| (B) TARGET | (B1) Hsu and Liu | 1990 | Standardized CPUE of albacore in the Indian Ocean caught by longline fisheries | Area(spp) Effect | Fishing area (targeting) effect is included in the GLM | Simple & easy to do | Can not specify targeting and Classification | Hsu and Liu (1990) |
| | (B2) Lee and Liu | 2000 | Standardized CPUE for yellowfin tuna caught by the Taiwanese longline fishery in the Indian Ocean | Non-target (bycatch) CPUE | Non-target CPUE is used in the GLM. No separation techniques were applied. CPUE of ALB and BET were used in GLM | CPUE of ALB and BET were used in the GLM as independent (input) variables for the by-catch effects, | YFT is the by-catch species for both Taiwanese regular LL and deep LL. So far, it is not easy to identify and separate those Taiwanese YFT target LL | Lee and Liu (2000) |
| | (B3) Chang (WPB, 2003) | 2004 | Catch Rate Analysis of the Indian Ocean Swordfish from Taiwanese Longline Fishery Using Generalized Linear Model | CPUE (GLM) | Apply quartile of catch ratio of studied species to the four main species (ALB BET YFT SWO) to GLM on set by set logbook data. | Simple and intuitive, could be incorporated into GLM model. For swordfish, the results very similar to using HBF information | Accuracy depends on quality of data. Data with few species reported that affect species comp need to be removed beforehand. | Chang et al or WPB (2003) |
| (C) OTHERS | (C1) Wang et al | 2002 | Using Fuzzy Synthesis Approach to Extract Fishing Efforts Directed on Albacore for Taiwanese Longline Fleets in the Indian Ocean | Fuzzy | No separation methods applied. Fuzzy spatial clustering classification | This method can avoid separating the fishing efforts directed to different species | No application yet ? Accuracy depends on the return rate and composition of return logbooks. | Wang et al (2002) |
| | (C2) Hsu | (idea) | | GAM | GAM including a proportion positive sets with a binomial model and positive catch rates with a delta lognormal error structure | Better than the catch ratio and Fuzzy approaches | Performance is unknown as this method has not yet applied to the real data. | |

## 3. Discussion

### (1) Goal

Our goal is to search approaches to conduct more accurate CPUE standardization or estimate abundance indices by GLM and other related methods reflecting real dynamic of Taiwanese LL fisheries when we don't have number of hook per basket information. We need to reflect real situation of fisheries. As real fisheries can be represented by species compositions.

### (2) Problems in the current methods

As the type (A) and (B) represent only part of species compositions, they can not always represent the real fisheries. In addition there are statistical problems in some cases, i.e., violation of acumination as in these types there are common variables used on both dependent and independent variables such as CPUE(BET) etc.

Even for the ideal situation using several classes of number of hooks per basket when such information are available, there are problems when there are strong 'Fakare' or 'LL floating due to underwater currents' because the expected target or gear configuration can not be represented any more in the data. Even such case, as long as we will understand the situation of fisheries by looking at species compositions.

For other types, the Fuzzy approach can avoid separating the fishing efforts directed to different species, bur accuracy depends on the return rate and composition of return logbooks. In addition, this approach can not reflect actual species compositions as it mask the species compositions by the fuzzy membership functions.

## 4. Prospects

Although we have pessimistic current situation, we have two potentially effective methods:

### (1) Species composition approach

As we need to reflect real fisheries in the GLM or other related method and the species composition are consider to reflect real fisheries. Thus it might be potentially effective method if we use species compositions (SPC), for example, in the GLM as follows (in case we have 4 major species in catch):

$$
\begin{aligned}
ln\ (CPUE\ Spp1) = &\ (mean) + Y + M + A \\
&+ (ENV\ or\ q\ related\ factors) + (SPC2 + SPC3 + SPC4) + errors
\end{aligned}
$$

Species composition SC1 is not included because (a) to avoid the potential statistical problem to use same variable in both independent and dependent variables and (b) SPC1 information is reflected by SPC 1-3 in the GLM as it is SPC= 1 – (SPC1-3).  This is a sort of the combined approach of type (A) because it takes account of LL configuration and target species.

### (2) GAM

This is the second approach in Other type, which was suggested by Dr Hsu, i.e., GAM including a proportion positive sets with a binomial model and positive catch rates with a delta lognormal error structure will be modeled. This is better than the catch ratio, type (A) and Fuzzy approaches. However, the performance is unknown as  this method has not yet applied to the real data.

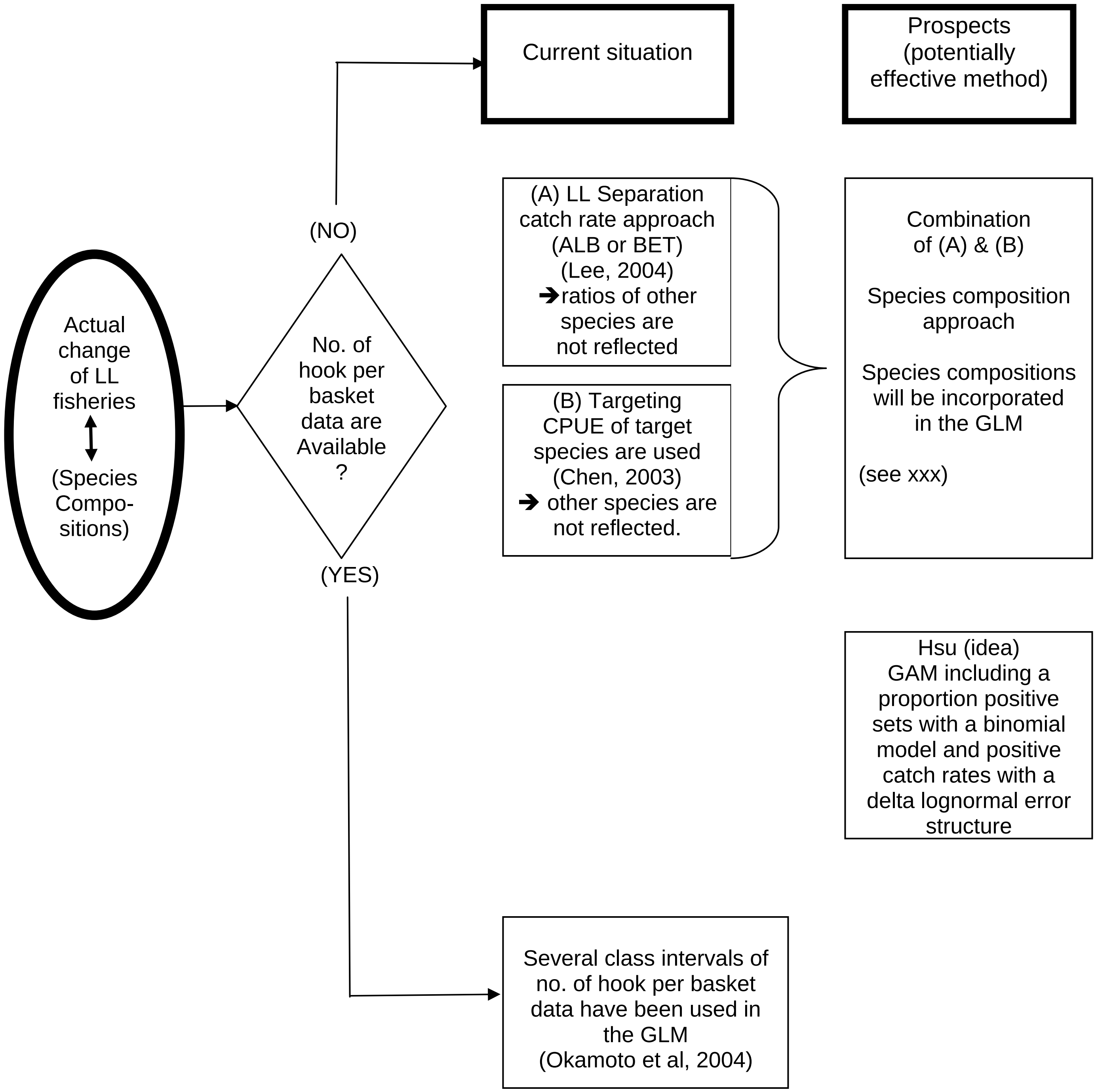
### (3)  Summary and Future work

The summary of our investigation is provide in Fig. 2. Based on our investigation (Fig. 2), we have some potential approaches to solve more accurate CPUE standardizing of the Taiwanese LL fisheries when we don’t have number of hooks per basket. Besides approaches discussed in this paper, we might have more ideas. Thus as a next step, we need to attempt and evaluate these approaches in the near future. Hence it is proposed to have a small working group to do this task.

## References

To be provided upon request.

Fig.2 (summary) How situation of the actual dynamic of LL fisheries can be reflected in GLM or other related methods : Current situation and Prospects



Appendix (A) Description of the approach type (A) : LL separation

(A1) Chang et al (1993)

*An Alternative Procedure to Segregate Mixed Longline Catch Data*

*Shui-Kai Chang, Chien-Chung Hsu and His-Chiang Liu*

*Misleading estimates of fish abundance would occur if the fishing efforts deployed are not on a specific species. Two kinds of fishing patterns are involved in Taiwanese longline fishers in the Indian Ocean, i.e., regular longline (RL) and deep longline (DL), which have quite different target species. Their catch and effort data therefore needed to be segregated so that the CPUEs would be valid and the consequent abundance estimation could be appropriately performed.*

*Two valuable segregating procedures have been developed, one by Suzuki and his associates, and the other one by Tuna Research Center of National Taiwan University. However, due to the difficulties to satisfy their demands on data, these two procedures can not be applied to the Taiwanese historical longline catches. An alternative procedure, which take advantages of the classification features of correspondence analysis and disjoint cluster analysis, was therefore suggested here to classify the mixed catches into RL catches and DL catches.*

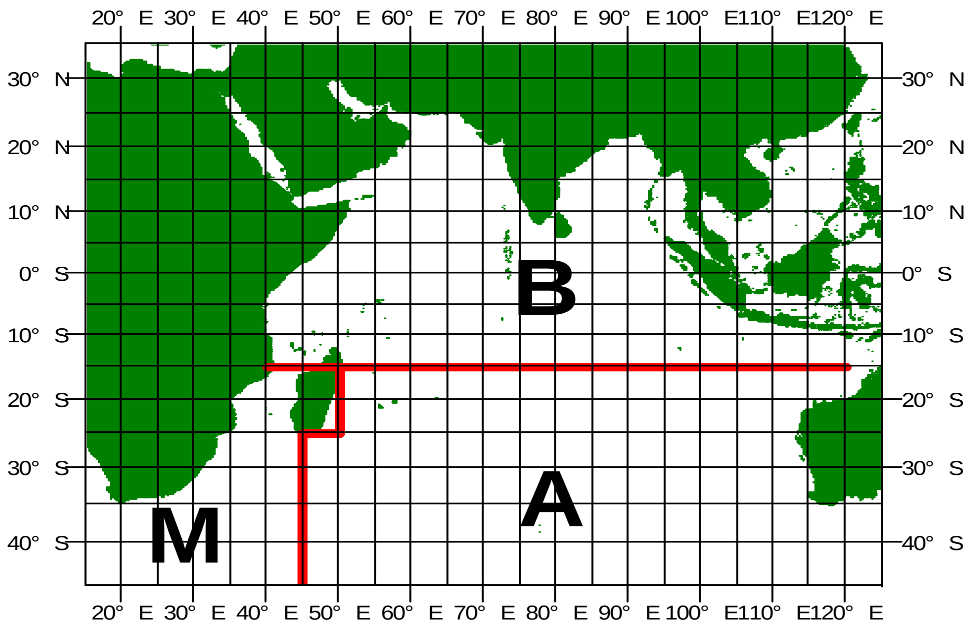
*Intuitively, the two new classified clusters (namely, SRL and SDL) were well segregated, and were rather coincident with the species composition and effort distribution of their fisheries. For the SRL clusters, albacore catches (target species of RL) composed more than 80% of the total tuna catches in the clusters; and for the SDL clusters, bigeye and yellowfin tunas together (target species of DL) composed more than 80%. The efforts made by SRL were mostly distributed southward of 10°S, where the spawning and immature albacore stock are concentrated, and those by SDL were occurred northward of 15°S, where the tropical tuna species distributed. A great discrepancy between their CPUE trends were also found in the segregation results.*

This paper is the first one deal with Taiwanese targeting issue and has persuaded people not to use the so-called 'registered type' as targeting index for Taiwanese longline fishery which was very common usage before that time. But, this method was not applied to GLM yet. However there are similar approached published. Basically, the method uses correspondence analysis to statistically decide how many groups could be separated in the data and then uses disjoint cluster analysis to separate the groups. We could apply GLM on the two separate groups. We could also apply GLM on the combined series with grouping result as an index in the model. It might be worthwhile to apply this method to the set by set logbook data. This method is better than a simple catch ratio method since it can reflect and examine the change in fishery.

(A2) Lin (1998)

The first attempt for the species (ALB) index. But no separation was made.

**A=ALB area (South of 15 S), B=BET area (North of 15 S); M= Mix area (Southwestern Indian Ocean)**




ALB ratio = ALB/(ALB+BET)

high ALB ratios in A(South of 15$^{o}$ S)

low ALB ratios in B(North of 15$^{o}$ S)

moderate ALB ratios in M(Southwest)

However, they had seasonal variations. The main result :

(1) ALB area had high ratios, i.e., 0.97 in Nov-May, 0.89 in Jun, ..., 0.82 in Sep-Oct.

(2) BET area had low ratios, i.e., 0.02 in Jan-Dec.

(3) Mix area had moderate or complex ratios, i.e., 0.96 in Nov-Dec, 0.84 in Mar-May, 0.51 in Jun-Jul, and 0.66 in Aug-Oct for the data with NHB = 9 and 10. As well as 0.04 in May-Sep for the data with NHB >= 12.

(A2a) Hsu and Liu (2000)

The current status of bigeye tuna in the Indian Ocean by a stochastic ASPM based on longline fishery data. In the CPUE standardization, Lin's criteria was applied to separate the deep LL.

(A2b) Hsu, Lee, Liu, and Liu (2001)

C. C. Hsu, H. H. Lee, Y. M. Liu, and H. C. Liu (2001) also adopted the Lin (1998) method and applied to the Indian bigeye tuna.

*Bigeye tuna, Thunnus obesus, is a valuable species distributing in tropical and temperate waters around the world. Taiwan is one of the leading nation fishing bigeye tuna in the three Oceans. In 1999, the catch of bigeye tuna amounted to 38,000 mt by Taiwanese longline fleets in the Indian Ocean. In this study, daily logbooks with set by set catch information were used, a partitioning of fishing effort made by different fishing types, say deep and regular types, was pursued before some discussion of spatial and catch composition distributions of different hooks used between floats. Then the new catch information was used to standardize catch per unit effort by general linear model, applied as the estimated abundance index of bigeye tuna for Taiwanese longline fishery in the Indian Ocean. The results showed that most years of the standardized time series trend were similar with Japanese trend, however, the trend after 1991 was opposite. This discrepancy needs to be in further investigated.*

(A3) Chen (1998)

Application of Lin's method. Chen (1998) adopted the Lin (1998) method and applied to the Indian albacore stock assessment.

Table 1 : Criteria (ratio of albacore tuna to the sum of albacore and bigeye catch)for demonstrating the albacore CPUE of regular longliners in the Indian Ocean.

| Month | Area A | Area B | Area M |
|---|---|---|---|
| Jan. | >=0.97 | >=0.02 | >=0.96 |
| Feb. | >=0.97 | >=0.02 | >=0.96 |
| Mar. | >=0.97 | >=0.02 | >=0.84 |
| Apr. | >=0.97 | >=0.02 | >=0.84 |
| May | >=0.97 | >=0.02 | >=0.84 |
| Jun. | >=0.89 | >=0.02 | >=0.51 |
| Jul. | >=0.95 | >=0.02 | >=0.51 |
| Aug. | >=0.91 | >=0.02 | >=0.66 |
| Sep. | >=0.82 | >=0.02 | >=0.66 |
| Oct. | >=0.82 | >=0.02 | >=0.66 |
| Nov. | >=0.97 | >=0.02 | >=0.96 |
| Dec. | >=0.97 | >=0.02 | >=0.96 |

(A4) Lee et al (2004)

Separation of the Taiwanese regular and deep tuna longliners
in the Indian Ocean using bigeye tuna catch ratios

Taiwanese longline fisheries (LL) in the Indian Ocean usually catch albacore tuna (ALB), swordfish (SWO) and yellowfin tuna (YFT) by the regular LL, on the contrary, bigeye tuna (BET) by deep LL. Thus these two types of LL are considered to be different gears as they catch different tuna species. Regular or deep type LL is defined by number hooks per basket (NHB), i.e., regular LL if $6 \leq$ NHB $\leq 10$ and deep LL if $11 \leq$ NHB $\leq 20$. However, the NHB information was available only in some of the recent LL data (1995-99). This situation has been causing problems of biased results in the stock analyses in the past. Under such backgrounds, the objective of our study is to explore an effective method to separate two types of LL considering species compositions. After various attempts, we found that some intervals of BET catch ratios were resulted to be most effective in separating regular and deep type LL, i.e., $0.8 \leq$ BET/(BET+ALB) $\leq 1$ and $0 \leq$ BET/(BET+ALB+SWO) $\leq 0.40$ respectively. Using these two separators, we classified the LL type known data set (1995-99) (learning data set). Then we found that 67.7% data were correctly classified, while 23.1% were un-classified (11.9 % for zero catches and 11.2% classified into both LL types) and 9.2% for mis-classification. Then, using the developed methods, we classified the unknown LL type in the historical data (1979-99) and computed nominal CPUE of four species. As a result, their CPUE trends are reasonably depicted.

Appendix (B) Description of the approach type (B): Targeting

(B1)Hsu and Liu (1990) not available

(B2) Lee and Liu (2000) not available

(B3) Chang (2003) or WPB (2003)

Catch Rate Analysis of the Indian Ocean Swordfish
from Taiwanese Longline Fishery Using Generalized Linear Model

*The target-shifting practice may affect the swordfish catch rate and hence needs to be accounted for in the model. Due to insufficient information on gear configuration (e.g., HPB), this study used three indices to express the target effect: (1) GLM-1: quartile of catch rates of the three other major species (albacore, bigeye and yellowfin tunas); (2) GLM-2: ranking of catch composition of the three other major species; and, (3) GLM-3: quartile of catch composition of swordfish. In the second index, proportion of the three major species were calculated in percentage for monthly 5x5 blocks and attached ranking for them from the higher percentage to lower. Those ranking were then categorized into four classes.*

*Although the standardized catch rates from the three GLMs showed similar trends, the GLM-1 and GLM-2 can only account for 30% of the variance, while GLM-3 can account for 80% of the variance. In GLM-3, the target factor has the highest F value, suggesting that this factor accounted for the highest amount of residual deviance. Even in GLM-1 and GLM-2, the combined target factor (Ualb+Ubet+Uyft or Ralb+Rbet+Ryft) also have significant high F-value. This phenomenon is different from the case in Campbell and Dowling (2002) when conducting standardization on Japanese bigeye fishery data since that Japanese fishery did not change their target as Taiwanese fishery did.*

*From the above results, it was worthwhile to conduct further analysis using the proportion of swordfish catch as an index for target factor. Since monthly aggregated data will lose important information on catch composition of a vessel in an operating set, therefore it will be preferable to conduct the analysis using set by set logbook data otherwise the catch rate trend will be misleading. However this practice will have to omit the information during 1967-1980 due to no logbook data available.*

*In reviewing the logbook data, it was noted that many fishermen did not report catches of all the 14 species on daily logbook records. Some records only reported catches of 1 or 2 species which will result in bias when calculating swordfish catch composition. In this case, the resulted residual distribution of the GLM fitting was deviant from normal. A new data set was created by excluded those records with number of species reported less than 4, and the residual distribution of the GLM fitting was improved and closer to normal. Fig. 8 shows its standardized catch rate, which has been adjusted by equation (2), with quartile of swordfish proportion in the catch as the target factor.*

*As mentioned in previous section, there is HPB information in 1995-2001 data. The information can be used to be a rough estimate of depth of hook for evaluating longline gear capture efficiency, or a better proxy index for target effect in the GLM. A GLM fitting taking HPB information as target factor was also performed and shows in Fig. 8 for comparison. It demonstrated the above standardized catch rate using swordfish proportion as target factor has correspondent trend with using HPB information, except for the first and last year which might be owing to incompleteness of HPB data.*

*The trend in Fig. 8 is quite different from that in Fig. 6. This was mainly caused by different effect in treating the target factor. Apparently the target factor has dominated the trend in Fig. 6 which follows the increasing trend of swordfish catch in 1990s (Fig. 1) due to increasing of targeting operations. The trend in Fig. 8 however did not show such pattern and was more consistent with the decreasing trend of Japanese CPUE analysis (Yokawa and Shono, 2000).*

*Target effect is important but complicate in deriving a reliable catch rate for Taiwanese catch data. Except the current approaches used in this study, there are some other options to deal with this issue, such as to define a specific categories for the swordfish catch proportion from the available HPB information and scientific observer program, or to use a more comprehensive statistical methods such as cluster analysis (Chang et al., 1993; He et al., 1997).*

Appendix (C) Description of the approach type (C) : Others

(C1) Wang, Hsu and Liu (2002) Fuzzy

*Using Fuzzy Synthesis Approach to Extract Fishing Efforts Directed on Albacore for Taiwanese Longline Fleets in the Indian Ocean*

*Indian albacore fishery is one of the most important tuna fisheries for Taiwanese longline fleets. The assessment of the Indian albacore stock is usually based on fishery-dependent data submitted from Taiwanese longline vessels. Moreover, those fishery data may contain two fishing types that are able to make standardizing catch per unit effort difficult. Therefore, in the present study, an alternative approach of fuzzy synthesis clustering is used to partition the fishing efforts from different fishing types, and the daily set catch information of logbooks from 1979 to 1997 is used as the fundamental data for this purpose. A fuzzy transformation is composed of weighting vector and membership function, in which the weighting vector used an unequal crisp value and the membership function used the distribution of percent catch of albacore in total of albacore, bigeye tuna, and yellowfin tuna under the factors of vessels' tonnage categories, fishing area, the number of hooks used and sea surface temperature. Subsequently, the result is obtained from the computation of fuzzy transformation, then, new catch, fishing effort and catch per unit effort series were obtained. The fuzzy synthesis is evidenced as one of the methods using for partitioning fishing efforts from different fishing types in preliminary.*

(C2) Hsu (suggestion) not available

## APPENDIX (D) REVIEWS AND PROSPECTS ON THE TREATMENTS OF CLASSIFICATIONS AND TARGETING FOR THE TAIWANESE LONGLINERS IN CPUE STANDARDIZATIONS - A REPORT OF THE SMALL GROUP MEETING

Report of the small group meeting
(9-11 AM, July 17, 2004 at the IOTC meeting room)

Participants
Anganuzzi, Nishida, Okamoto, Shono, Chang, Mosqueira and Fujiwara

Topic
Reviews and prospects on the treatments of classifications and targeting for the Taiwanese longliners in CPUE standardizations
(IOTC/WPTT-04-10)

Based on detail and extensive discussions, we consider that following two approaches may be effective to reflect real LL fisheries in CPUE standardization when we don't have number of hook between basket information. Japanese and other scientists interesting in these approaches will attempt for BET using the Japanese LL data from 1970-85 in the tropical and/or the whole Indian Ocean. The Japanese LL data will be used as they have a longer period and wider fishing grounds as test (learning) data sets in Taiwan. Then results can be evaluated by comparing with those with real information (learning data sets). Results plan to be reported in the next WPTT in 2005.

It was recognized that for SWO, the approach developed in the last WPB is effective and useful as the results were very similar to those of the learning data sets (data with the number of hook per basket).

(1) Trip based approach

LL trip (longer term) instead of short (e.g., set-by-set) term based analyses is proposed to investigate the target practices. This approach likely more reflects real targeting practices as the data are aggregated into one trip which usually uses consistent (common) gear configurations. Based on this approach, we can learn the relationship between gear configuration and targeting practices. Hence, once we learn their characteristics we can classify for example, Regular, Deep or Super Deep LL, which is likely more robust than the catch rate approach by Lee et al (2004) which include heterogeneous gear configurations creating biases (misclassification). Then, Using the hook per basket known data (leaning data set), we can evaluate accuracy (performance) of this approach.

(2) Species composition approach

We need to reflect real changes of LL fisheries in the GLM or other related methods in CPUE standardization. As the species compositions are considered to reflect real changes of LL fisheries, CPUE standardization incorporating species compositions might be potentially effective and useful method. For example, if we want to standardize BET and if we have 4 major species in catch (BET, ALB, YFT and SWO), we will have the following GLM model:

$$\ln(\mathit{CPUE\ BET}) = (\mathit{mean}) + Y + M + A + (\mathit{ENV\ or\ q\ related\ factors}) + (ALB' + YFT' + SWO') + \mathit{errors}$$

where, ALB', YFT' and SWO' are the 10 categories of species compositions (as example). Then, if we have 22%(BET), 39%(YFT), 3%(SWO) and 36%(ALB), YFT'=3, SWO'=0 and ALB'=3. BET species composition is not included because (a) to avoid the potential statistical problem to use same variable in both independent and dependent variables and (b) BET' information is reflected by other species compositions in the GLM as BET'= 1 – (YFT'+ALB'+SWO'). This is a sort of the combined approach of current two types (separation and targeting based) approaches.