INTER-AMERICAN TROPICAL TUNA COMMISSION SCIENTIFIC ADVISORY COMMITTEE FOURTH MEETING

La Jolla, California (USA)

29 April - 3 May 2012

DOCUMENT SAC-04-05B

ANALYSES OF JAPANESE LONGLINE OPERATIONAL-LEVEL CATCH AND EFFORT DATA FOR BIGEYE TUNA IN THE EASTERN PACIFIC OCEAN

Cleridy E. Lennert-Cody, Hiroaki Okamoto and Mark N. Maunder

SUMMARY

Analyses of operational-level longline data for bigeye tuna in the eastern Pacific Ocean were conducted to study the effects of differences in fishing efficiency among vessels on the estimated long-term trend in the index of relative abundance. Results suggest that when differences in fishing efficiency among vessels are taken into consideration, the long-term trend in the index is slightly more pessimistic. This result is consistent with findings of similar analyses of data for bigeye from the western Pacific Ocean. Inspection of residuals from the standardization models suggests that incorporation of environmental variables, and terms for interactions between the number of hooks between floats and location and environment, may help to improve model fit.

1. BACKGROUND

Trends in longline catch-per-unit-effort (CPUE) are a very important driver in the bigeye and yellowfin tuna assessment of the eastern Pacific Ocean (EPO) (Aires-da-Silva and Maunder 2012, and references therein). In recent years concerns about possible trends in catchability due to improved vessel performance have been raised. Recent analyses of Japanese operational-level longline data from the western Pacific Ocean (Hoyle *et al.* 2010; Hoyle and Okamoto 2011) have identified differences in fishing efficiency among vessels previously not accounted for in the generalized linear models (GLMs) used for estimation of indices of relative abundance. To explore the extent to which similar changes in vessel efficiency also may be occurring in the EPO, an exploratory analysis of Japanese operational-level longline data for bigeye from the EPO was conducted and is presented herein. These analyses were patterned after studies of data from the western Pacific Ocean (Hoyle *et al.* 2010; Hoyle and Okamoto 2011) to allow for comparability.

2. DATA AND DATA PREPARATION

By-set Japanese longline data for 1975-2011 were used in these analyses to be consistent with the current time period for the EPO bigeye stock assessment (Aires-da-Silva and Maunder, 2012). The data were originally structured in two data files, one for 1975-1993 and one for 1994-2011. Each record in the data files represented an individual longline set. The data fields were: year, month and day of the set; the latitude and longitude of the set (1° resolution); the Japanese vessel name; the call sign of the vessel; the start date of the trip; the tonnage of the vessel; the fishing category of the vessel (offshore and distantwater); the license number of the vessel; the target species group of the set (swordfish, shark, other including tunas); the longline characteristics (mainline material, branchline material, bait type, number of hooks between floats; number of hooks on the longline); catch in numbers of each tuna species (albacore, bigeye, yellowfin). Not all variables were available throughout the entire time period. In particular, call

sign information was only available beginning in 1979, bait information was only available for 1975-1993, and longline material information was only available beginning in 1994.

These Japanese longline operational data were used in this study under the following conditions:

- 1. The usage of the data is strictly limited to the purpose of this collaborative work;
- 2. The data can be used only during this collaborative work;
- 3. The participant can use the data only on the personal computer prepared by Japanese scientists of NRIFSF, and any copy of the data from the computer is not permitted; and
- 4. Any document or presentation derived from the result of this collaborating work should be consulted beforehand to the Japanese Fishery Agency and NRIFSF scientists.

Data processing included recoding of missing values, removing records with unusual data, and creating several derived variables. All data processing was done in R (R Development Core Team, 2012). Missing values were re-coded from "." (or 0 for trip start date) to NA for the following variables: number of hooks between floats, tonnage of the vessel, vessel call sign, vessel license, and trip start date. The fishing category of the vessel was re-coded for the 1975-1993 period to be consistent with the code values used in 1994-2011. Records with unusual values in several variables were identified and deleted: day of the month greater than 31, number of hooks between floats less than 3 or greater than 30, and total number of hooks on the longline greater than 6000. Duplicate records for the same set were removed. Duplicate records were identified as those records with the same date and location of fishing for the same vessel call sign and Japanese vessel name and year. For the 1975-1993 period there were 1,118 such records out of 895,170 (0.12%), and 65 records out of 645,859 (0.01%) for the 1994-2011 period.

The data set also was trimmed prior to analysis to exclude: records with missing data in key fields (call sign, vessel name, trip start date, number of hooks between floats and total number of hooks), records for sets outside the EPO (longitude west of 150°W, latitude north of 40°N or south of 40°S), and records from vessels that were not distant-water vessels (vessels other than distant-water vessels were likely research and training vessels). This limited the data set to the time period 1979-2011, primarily because vessel call sign was not available before 1979. The data processing reduced the number of total records from 1,748,019 to 1,424,871. For this trimmed data set, the following new variables were created: year-quarter, 5° area indicator.

Several data summaries and analyses reference the most recent IATTC stock assessment areas for bigeye, which are shown in Figure 1.

3. SUMMARIES OF GEAR, FLEET AND CATCH CHARACTERISTICS

The number of hooks between floats showed considerable temporal and spatial structure (Figures 2-4). The median number of hooks between floats increased from 10 to about 15 between 1979 and 1994, and has increased slightly to 17 in recent years (Figure 2). Since 1994, when information on longline material was first available, the number of hooks between floats has been less variable among sets with nylon mainline and branchlines (which is the dominant material combination) than for sets with mainlines and branchlines of "other" materials (Figure 3). Fewer hooks between floats were used in the more temperate areas of the EPO as compared to the more tropical equatorial regions (Figure 4), and thus the temporal changes in hooks between floats have varied spatially across the EPO.

The median number of hooks per longline set has also varied somewhat through time, generally increasing over the period 1979-2011 (Figure 5). In addition, the variability in number of hooks per longline set about the median shows temporal structure.

The time period of data representing each unique value of call sign (assumed to uniquely identify individual vessels) is shown in Figure 6. There were 1,265 unique values of call sign for the 1979-2011 period. Call sign values in Figure 6 are ordered according to the date of the most recent data in the data

set and then the date of the earliest data in the data set, and thus, at the bottom of the figure is shown the reporting history for the vessels representing the recent fishery.

Spatial-temporal changes in catch composition within the EPO were clearly visible over a wide range of scales for the three catch species (bigeye, yellowfin, and albacore; Figures 7-9). In general, the vast majority of sets in most regions of the EPO caught at least some amount of bigeye. However, overall decreases in CPUE can be seen across the fishery for both bigeye and yellowfin for the last decade as compared to the fishery in the 1980s and 1990s. As compared to the rather spatially consistent annual pattern for albacore, spatial pattern in catch rates for bigeye and yellowfin varied from year to year and were often contrasting, with areas of higher catch rates of bigeye being areas of lower catch rates for yellowfin.

The frequency distribution of bigeye catch per set (in numbers) is skewed towards large values, with only a few percent of the sets having no catch (Figure 10). The frequency distribution of bigeye caught per hook (Figure 11) suggests occasional rounding of numbers of fish and numbers of hooks, but has a similar overall shape to that of catch per set.

4. ANALYSIS OF TRENDS IN BIGEYE TUNA CPUE, WITH AND WITHOUT CALL SIGN EFFECT

Analyses of the effect of differences in fishing efficiency among vessels on the estimates of bigeye relative abundance indices were conducted separately for each of the four IATTC stock assessment areas (Figure 1). For practical reasons, the analyses were limited to the data of vessel call signs that were well represented: within each area, the data were limited to those call signs within the upper 25% of sets per call sign. The amount of data available by area is shown in Table 1. Given the shape of the overall frequency distribution of bigeye catches (Figures 10-11), negative binomial models (bigeye counts) and lognormal models (bigeye CPUE) were used for these data.

The following linear ("LM"; equations (1)-(2) below) and generalized linear (negative binomial, "NB"; equations (3)-(4) below) models were fitted to the data by area:

- 1. $\log (CPUE + 0.0002) = \text{constant} + \text{year-quarter effect} + 5^{\circ} \text{ area effect} + f(\text{hooks between floats})$
- 2. $\log (CPUE + 0.0002) = \text{constant} + \text{year-quarter effect} + 5^{\circ} \text{ area effect} + f(\text{hooks between floats}) + \text{call sign effect}$
- 3. log (μ) = constant + β •log(number of hooks) + year-quarter effect + 5° area effect + *f*(hooks between floats)
- 4. $\log (\mu) = \text{constant} + \beta \cdot \log(\text{number of hooks}) + \text{year-quarter effect} + 5^\circ \text{ area effect} + f(\text{hooks between floats}) + \text{call sign effect}$

where μ is the mean bigeye catch (count), β the slope corresponding to the linear term log(number of hooks), and *f* represents a natural spline smooth of degree 6. The LM model was not fitted to the data of Area 1 because of the large percentage of sets with 0-valued bigeye catches (Table 1). For comparison to the NB results, a Poisson GLM was also fitted to the data of Area 2. The form of the models above was selected to be consistent with analyses for the western Pacific Ocean (Hoyle *et al.* 2010). To provide more information on the relationship between hooks and catch, log(number of hooks) was included in the model as a linear term, not as an offset, thereby obtaining an estimate of the slope coefficient (rather than assuming a value of 1.0). In addition, in some areas (see Table 2) model run time for the NB GLM was extremely slow when estimating the scale parameter (θ) for the model with a call sign effect. In such cases, the model was fitted with the value of θ from the model without a call sign effect. Future work could focus on improvements to these models (see also below).

Based on the Akaike Information Criterion (AIC), adding a call sign effect to the models improved model fit for all of the models considered (Table 3). The year-quarter effect coefficients from the NB GLMs, by area, are shown in Figures 12-15. In Area 1, there was little difference in year-quarter effect coefficients

from models with and without call sign effects. For Areas 2-4, the year-quarter effect coefficients from models with a call sign effect were slightly smaller than those from models without a call sign effect, beginning in the early 1990s through 2011.

Differences between year-quarter effect coefficients of models with and without a call sign effect, by area, were summarized as the ratio of mean-standardized year-quarter effect coefficients, r:

$$r_{i} = \frac{\frac{\underline{\gamma_{i_without}}}{\overline{\underline{\gamma}_{._without}}}}{\frac{\underline{\gamma_{i_with}}}{\overline{\overline{\gamma}_{._with}}}}$$

where γ_{i_*} is $exp(i^{th}$ year-quarter effect coefficient) (at i=1, the coefficient value was set to 0) and $\overline{\gamma}_{*}$ is the mean of the exponentiated coefficients (mean over *i*). The time series of *r* are summarized in Figure 16 for the NB GLM. By area, the time series were similar for the LM (not shown) and NB GLM (and Poisson GLM, not shown) models, with no long-term trend in Area 1 and long-term trends in Areas 2-4. In Areas 2-4, the time of transition from r < 1.0 to r > 1.0 varied somewhat by area, but in all three areas the transition occurred in the early to middle 1990s. The amplitude of the departure (+/-) of *r* from 1.0 also differed by area, with the smallest positive and negative amplitudes occurring in Area 3.

Standardized trends for bigeye from NB GLM models, with and without a call sign effect, were computed for Area 2 (Figure 17). Area 2 was selected because the NB GLM had the best fit to the data of that area (based on inspection of diagnostic plots, although improvements to the model could be made; see Figure 18) and Area 2 has been one of the dominant fishery areas within the EPO. The standardized trend from the model with a call sign effect suggests a slightly greater decreasing trend from 1979 to 2011 as compared to that from the model without a call sign effect. These standardized trends are based on partial dependence (Hastie *et al.* 2009), which summarizes the effect of year-quarter on catch amount, having accounted for the average effects of the other variables in the model.

5. DISCUSSION AND FUTURE WORK

In general, accounting for differences in fishing efficiency among vessels (inclusion of a call sign effect) in the estimation of year-quarter indices for the EPO (Figures 16-17) had a similar effect as that seen for the western Pacific Ocean (Hoyle *et al.* 2010; Hoyle and Okamoto 2011). In three of the four EPO stock assessment areas, the ratio of mean-standardized year-quarter coefficients showed smoothly varying trends, transitioning from increasing to decreasing fishing power around the mid-1990s. Similarly, for Area 2, the only stock assessment area of the EPO for which a standardized index was computed, the standardized trend based on a model with a call sign effect tended to be slightly more pessimistic than the trend from a model without a call sign effect. An analysis of residuals from the model fitted without a call sign effect indicated that for some call signs there was a temporal trend in the residuals over the time period fishing activity (Figure 19). Thus, if further work on fishing power is undertaken, the models might be improved by allowing for simple trends by call sign, instead of assuming that any change in fishing power through time was constant (as is done in the models described above).

Addressing spatial misfit in the standardization model may be, however, more important than modeling vessel-specific differences in fishing power. Even with a call sign effect, the fit of the NB GLM to the data indicated shortcomings (Figure 18). The same was true of the LM and Poisson models (not shown). This lack of fit may be due to a misspecification of the distributional aspect of the model or to a misspecification of the mean structure (or both). Annual maps of residuals from the NB GLM model with a call sign effect for Area 2 showed spatial structure which varied by year, perhaps indicating that the model for the mean structure might be improved by added environmental predictors (*e.g.*, SST and mixed layer depth), in addition to predictors capturing smaller-scale spatial differences in gear use (*e.g.*, interactions between the number of hooks between floats and location and environment) and indicators of

targeting. Spatial structure in the number of hooks between floats, on both large scales (Figure 4) and small scales (not shown), was evident in the operational-level data. Previous studies (*e.g.*, Maunder *et al.* 2006) have illustrated the importance of considering smaller-scale spatial (habitat) effects. Negatively-correlated spatial structure among catch species was visible over several scales in the data of the present study, although it is not clear that this structure relates to spatial patterns in targeting. Trends in targeting are also being studied in the western Pacific Ocean longline fishery (Hoyle *et al.* 2010; Hoyle and Okamoto 2011). Addressing these issues may require specifying the standardization models as generalized additive models instead of GLMs.

ACKNOWLEDGMENTS

Special thanks to the NRIFSF for granting access to the operational-level data, to our Japanese colleagues for their assistance and hospitality in Shimizu, and to Simon Hoyle for helpful discussions.

REFERENCES

- Aires-da-Silva, A. and Maunder, M.N. 2012. Status of bigeye tuna in the eastern Pacific Ocean in 2011 and outlook for the future. In: IATTC Stock Assessment Report 13. http://www.iattc.org/PDFFiles2/StockAssessmentReports/SAR-13-BETENG.pdf
- Hastie, T., Tibshirani, R., Friedman, J. 2009. The Elements of Statistical Learning, Second Edition. Springer.
- Hoyle, S. D., Shono, H., Okamoto, H. and Langley, A.D. 2010. Analysis of Japanese longline operational catch and effort for bigeye tuna in the WCPO. Document WCPFC-SC6-2010/SA-WP-02.
- Hoyle, S.D. and Okamoto, H. 2011. Analyses of Japanese longline operational catch and effort for bigeye and yellowfin tuna in the WCPO. Document WCPFC-SC7-2011/SA IP-01.
- Maunder, M.N., Hinton, M.G., Bigelow, K.A. and Langley, A.D. 2006. Developing indices of abundance using habitat data in a statistical framework. Bulletin of Marine Science 79:545-559.
- R Development Core Team 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/

TABLE 1. Information of the data sets used in the LM and NB GLM (Poisson GLM) analyses.**TABLA 1.** Información de los conjuntos de datos usados en los análisis ML y MLG NB (Poisson MLG).

	# records used	% records	% BET catch	% records with	# unique	Minimum
	for analysis	retained	retained	no BET catch	vessel call	number of
					signs	records per
						call sign
	# registros	% registros	% captura de	% registros sin	# señales de	Número míni-
	usados para	retenidos	patudo	captura de	llamada únicas	mo de regis-
	análisis		retenido	patudo		tros por señal
						de llamada
Area 1	89277	72.3	74.0	8.8	187	211
Area 2	196096	66.7	65.1	2.4	274	367
Area 3	486975	72.5	63.0	0.9	268	1031
Area 4	171386	66.5	67.8	3.1	194	429

TABLE 2. Estimates of NB scale parameter (θ) and of the coefficient for log(number of hooks), from the NB GLMs, with and without call sign effect. Dashes indicate estimate not obtained.

TABLA 2. Estimaciones del parámetro de escala NB (θ) y del coeficiente de log(número de anzuelos), de los MLG NB, con y sin efecto de señal de llamada. Los trazos indican estimación no obtenida.

	θ , without call	θ , with call sign	log(hooks) coefficient,	log(hooks) coefficient,	
	sign effect	effect	without call sign effect	with call sign effect	
	θ , sin efecto de	θ , con efecto de	coeficiente	coeficiente	
	señal de llamada	señal de llamada	log(anzuelos), sin efec-	log(anzuelos), con efecto	
			to de señal de llamada	de señal de llamada	
Area 1	2.348831		0.4198282	0.4736601	
Area 2	1.70188	1.747699	0.1455818	0.1081811	
Area 3	2.974777		0.193419	0.2679528	
Area 4	1.84593	1.920904	0.2025275	0.2473667	

TABLE 3. AIC values and difference in AIC for models with and without call sign effect. **TABLA 3.** Valores de AIC y diferencia en AIC para modelos con y sin efecto de señal de llamada.

	AIC, without call	AIC, with call sign	Difference in AIC
	sign effect	effect	(without – with)
	AIC, sin efecto de	AIC, con efecto de	Diferencia en AIC
	señal de llamada	señal de llamada	$(\sin - \cos)$
LM—ML			
Area 2	545059.2	541545.3	3513.9
Area 3	1073220	1062439	10781
Area 4	451519.5	445699.6	5819.9
NB GLM—MLG NB			
Area 1	674408.6	668326.5	6082.1
Area 2	1535008	1530663	4345
Area 3	3499973	3488086	11887
Area 4	1338387	1332510	5877



FIGURE 1. Map of the most recent IATTC stock assessment areas for bigeye tuna. **FIGURA 1.** Mapa de las áreas de evaluación del atún patudo más recientes de la CIAT.



FIGURE 2. Box-and-whisker plots of the number of hooks between floats by year for 1979-2011. **FIGURA 2**. Gráficas de caja y bigote del número de anzuelos entre flotadores por año durante 1979-2011.



FIGURE 3. Relationship between the number of hooks between floats and mainline and branchline materials since 1994. Pooled over years, the percentage of sets with the following mainline-branchline material combinations was: 77.4% for nylon-nylon (upper left panel); 3.7% for nylon-other material (upper right panel); 5.3% other material-nylon (lower left panel); 12.9% for other material-other material (lower right panel).

FIGURA 3. Relación entre el número de anzuelos entre flotadores y material del reinal y brazoladas desde 1994. Agrupados para todos los años, el porcentaje de lances con las siguientes combinaciones de material reinal-brazolada fue: 77,4% nylon-nylon (panel superior izquierdo); 3,7% nylon-otro material (panel superior derecho); 5,3% otro material-nylon (panel inferior izquierdo); 12,9% otro material-otro material (panel inferior derecho).



FIGURE 4a. Change in the number of hooks between floats by stock assessment area (Figure 1) and year, in millions of hooks.

FIGURA 4a. Cambio en el número de anzuelos entre flotadores por área de evaluación (Figura 1) y año, en millones de anzuelos.



FIGURE 4b. Change in the number of hooks between floats by stock assessment area and year, in percent hooks (*i.e.*, same as Figure 4a, but in percent).

FIGURA 4b. Cambio en el número de anzuelos entre flotadores por área de evaluación de poblaciones y año, en porcentaje de anzuelos (es decir, igual que la Figura 4a, pero en porcentaje).





FIGURA 5. Gráficas de caja y bigote del número de anzuelos por lance, por año, durante 1979-2011. La línea de trazos roja representa la mediana general del número de anzuelos por lance.



FIGURE 6. Date (year-quarter) of first and last records for each unique vessel call sign, ordered first by the date of the last record in the data set and then by the data of the first record in the data set. **FIGURA 6**. Fecha (año-trimestre) del primer y último registro de cada señal de llamada única de buque, ordenado primero red por la fecha del último registro en el conjunto de datos y luego por los datos del primer registro en el conjunto de datos.



FIGURE 7. Average bigeye tuna CPUE (number of fish per hook). **FIGURA 7**. CPUE media de atún patudo (número de peces por anzuelo).



FIGURE 8. Average yellowfin tuna CPUE (number of fish per hook). **FIGURA 8**. CPUE media de atún aleta amarilla (número de peces por anzuelo).



FIGURE 9. Average albacore tuna CPUE (number of fish per hook). **FIGURA 9**. CPUE media de atún albacora (número de peces por anzuelo).





FIGURA 10. Distribución de la captura por lance de atún patudo (número de peces) durante 1979-2011. No se incluye el 1% de observaciones con capturas de más de 100 peces.



FIGURE 11. Distribution of number of bigeye tuna caught per hook, 1979-2011. Not shown is the righthand tail at values between 0.05 and 5.0, which represents less than 5% of the observations. **FIGURA 11**. Distribución del número de atunes patudo capturado por anzuelo, 1979-2011. No se incluye la cola a la derecha de valores entre 0,05 y 5,0, que representa menos del 5% de las observaciones.



FIGURE 12. Estimated year-quarter effect coefficients from the NB GLM analysis for Area 1, for models with (red dashed line) and without (black line) call sign effect. Y-axis was cropped at a value of -4 in order to show detail between -1 and 0.

FIGURA 12. Coeficientes estimados del efecto año-trimestre del análisis MLG NB del Área 1, correspondientes a modelos con (línea de trazos roja) y sin (línea negra) efecto de señal de llamada. El eje y fue cortado en el valor de -4 a fin de ilustrar el detalle entre -1 y 0.





FIGURA 13. Coeficientes estimados del efecto año-trimestre del análisis MLG NB del Área 2, correspondientes a modelos con (línea de trazos roja) y sin (línea negra) efecto de señal de llamada.





FIGURA 14. Coeficientes estimados del efecto año-trimestre del análisis MLG NB del Área 3, correspondientes a modelos con (línea de trazos roja) y sin (línea negra) efecto de señal de llamada.





FIGURA 15. Coeficientes estimados del efecto año-trimestre del análisis MLG NB del Área 4, correspondientes a modelos con (línea de trazos roja) y sin (línea negra) efecto de señal de llamada.



FIGURE 16. Ratios of standardized year-quarter effect estimated coefficients, by area, from NB GLM models, with and without vessel call sign effects (see text for ratio description). Dashed red line is at a ratio value of 1.0.

FIGURA 16. Razones de coeficientes estimados del efecto año-trimestre estandarizado, por área, de los modelos MLG NB, con y sin efecto de señal de llamada de buque (ver descripción en el texto). La línea de trazos roja indica el valor de razón de 1,0.



FIGURE 17. Standardized index for bigeye tuna for Area 2 (y-axis is number of fish per hook). Red dashed line is from model with call sign effect, black line is from model without call sign effect; turquoise dashed line is at the average index value.

FIGURA 17. Índice estandarizado de atún patudo para el Área 2 (eje y es el número de peces por anzuelo). La línea de trazos roja es del modelo con efecto de señal de llamada, la línea negra del modelo sin efecto de señal de llamada; la línea de trazos turquesa indica el valor medio de los índices.



FIGURE 18. Diagnostic plots for the NB GLM with call sign effect for Area 2.

FIGURA 18. Gráficos diagnósticos correspondientes al MLG NB con efecto de señal de llamada para el Área 2.



Ordered can Sign-Serial de namada Orderiada

FIGURE 19. Boxplots of estimated slope coefficients (y-axis) *versus* call sign (x-axis; by groups of ~ 10 unique call sign values, ordered based on start and end dates in the data set). Estimated slope coefficients were obtained from a linear regression model fitted separately to the working residuals associated with each call sign (working residuals from the NB GLM for Area 2 without a call sign effect). Red dashed line indicates an estimated slope value of 0.

FIGURA 19. Gráficos de caja de coeficientes estimados de pendiente (eje y) como función de señal de llamada (eje x; por grupos de ~ 10 valores de señal de llamada única, ordenados de acuerdo a fecha de inicio y fin en el conjunto de datos). Se obtuvieron los coeficientes de pendiente de un modelo de regresión lineal ajustado por separado a los residuales asociados a cada señal de llamada (residuales del MLG NB para el Área 2 sin efecto de señal de llamada). La línea de trazos roja indica un valor de pendiente estimado de 0.