SCIENTIFIC COMMITTEE
EIGHTH REGULAR SESSION

7-15 August 2012
Busan, Republic of Korea

Decomposition of Tuna Length Frequency Data from Port Sampling in Papua New Guinea

WCPFC-SC8-2012/ SA-IP-06

Brian Kumasi, Thomas Usu, Leontine Baje and Ludwig Kumoru[1]

[1] Papua New Guinea National Fisheries Authority

Decomposition of tuna length frequency data from port sampling in Papua New Guinea

Brian Kumasi, Thomas Usu, Leontine Baje and Ludwig Kumoru

Papua New Guinea National Fisheries Authority

2012

Abstract

A histogram of frequencies of length often shows distinct modes that hypothetically represent distinct age classes. A variety of methods have been proposed for the analysis of length composition into age groups and particularly for the determination of the parameters of specified growth curves. National Fisheries Authority (NFA) of PNG has been collecting landings data through a port sampling program located at the major landing and trans-shipment ports in PNG. In this paper, length frequency analysis was performed to decompose the length frequency distribution of tuna obtained through port sampling for a 12 month period and generate length-at-age estimates and a length converted catch curve to determine mortality estimates for the tuna purse seine fishery in Papua New Guinea. The growth parameters for skipjack tuna (*Katsuwonus pelamis*) was estimated as L_{inf}=150cm, K=0.3 and t_0=0.14. The fishing mortality was estimated at F=-0.31. This work can be considered as the basis of providing an initial look at developing a more formal in-country analysis framework of PNG tuna catches and effort.

Background

The most common information collected in fisheries research is length, perhaps due to its ease of collection. A histogram of frequencies of length often shows distinct modes that hypothetically represent distinct age classes (Quinn and Deriso 1999). Length frequency analysis has been used to decompose a length frequency histogram into component age classes (Quinn and Deriso 1999), having evolved from the theory of mixtures of distributions (Hasselblad 1966) using maximum likelihood theory to decompose components from an overall distribution. A variety of methods have been proposed for the analysis of length composition into age groups and particularly for the determination of the parameters of specified growth curves (Shepherd 1987). Age composition is an increasingly important source of information for understanding fish population dynamics. By obtaining length, weight and age information it is possible to follow year-classes through time to understand growth, to monitor population changes over time and obtain age-specific estimates of population parameters (Quinn and Deriso 1999).

Catches of skipjack tuna (*Katsuwonus pelamis*) and yellowfin tuna (*Thunnus albacares*) in the archipelagic waters of PNG have steadily increased since 1997. Total tuna catches in PNG's EEZ have averaged about 250,000 tonnes per year over the past decade peaking at approximately 466,000 tonnes in 2007 (Kumoru 2008). Since 2008, the National Fisheries Authority (NFA) of PNG has been collecting landings data through a port sampling program located at the major landing and trans-shipment ports in PNG. Through the port sampling program, a wealth of length frequency data has been collected and preliminary analysis of this data set has been carried out (Kumasi, et al. 2010).

Length based methods are often used when length composition data for the total period is available for a one year period only. Catch curves and cohort analysis can be applied under these conditions as their basic assumption is the constant parameter system, whereby the picture presented by all length classes caught during one year reflects that of a single cohort during its entire life span (Sparre and Venema 1992).

In this paper we shall perform a length frequency analysis to decompose the length frequency distribution of tuna obtained through port sampling for a 12 month period and generate length-at-age estimates and a length converted catch curve to determine mortality estimates for the tuna purse seine fishery in Papua New Guinea.

Port Sampling Data

An estimated 20% by weight of all tuna catches that was taken from the ports of Lae, Madang, Wewak and Rabaul by fishing vessels to either land or trans-ship were sampled. A well in a vessel was divided into three layers (Top, Middle and Bottom) in which a number of nets were selected from each layer and were sampled. The number of nets per layer was selected or determined based on the total weight of catch in the well to ensure that 20% of the catch, in the particular storage well was sampled. This was done for all storage wells on a vessel that were unloaded. All fish in the net were identified to species level and fork length (cm) measurements were taken for all the fish in the net.

Length-at-age

The method of length frequency analysis utilises maximum likelihood theory to decompose components from an overall distribution (Shepherd 1987). The derivation of the maximum likelihood method has the following assumptions:

1. The length frequency distribution is multinomial:

$$f(L') = \binom{L}{L'_l \ldots L'_J} \prod_{l=1}^{J} p_l^{L'_l}$$

   The log likelihood is then:

$$\ln f(L') = C + \sum_{l=1}^{J} L'_l \ln p_l \ \ldots \qquad (1)$$

   where C is a constant.

2. The probability density function (PDF) of length $f_a(\boldsymbol{L})$ for each age group *a* is described by unknown parameter vector *υ*. Usually the normal PDF is used with mean μ_a and a standard deviation (SD) σ_a, given by,

$$fa(L) = \frac{1}{\sqrt{2}\,\sigma_a} exp\left[\frac{1}{2\sigma_a^2}(L - \sigma_a)\right] \quad \ldots \qquad (2)$$

   for the normal pdf is made up of the parameter vectors **μ** and **σ**. Alternatively, the log normal distribution can be used if the component distribution is thought to be positively skewed.

3. The number of age groups *A'* is known.

4. Each length measurement falls into one and only one interval *l*.

The basic tool for length-based methods is the conversion of length-based data aged-based data principally by use of the von Bertallanfy growth function (VBGF):

$$L_t = L_\infty * \left[1 - e^{(-K*(t-t_0)}\right] \ \ldots \qquad (3)$$

where the size of a fish at a particular time *Lt* is a function of the maximum size class L_∞, the growth rate parameter K, time when fish enter the system t_0.

The conversion of length into age is then done with the inverse VBGF:

$$t(L) = t_0 - \frac{1}{K}\ln\left(1 - \frac{L}{L_\infty}\right) \quad \ldots \qquad (4)$$

Length converted catch curve

The construction of a catch curve is the most common approach to estimate the total mortality of a cohort. Assuming constant recruitment and constant mortality, the length converted catch curves take the form:

$$ln\left(\frac{C_i}{dt}\right) = a + Zt_i' \qquad \ldots \qquad (5)$$

where, catch numbers for catch class *C,* the time needed for the fish to grow through the time class *i*, total mortality *Z* and a constant *a*. For a direct estimation of fishing mortality or total mortality, length based methods investigate the shape of the length-frequency composition. These methods are based on the assumption that large fish will be abundant under lower fishing mortality and rarer under high fishing mortality. Similarly length based methods assume steady-state recruitment and mortality, exponential decline of the population with natural mortality, and knife edge selection (Ziegler, Welsford and Constable 2011).

The effects of total mortality depend on fish growth parameters and results are often expressed as the ratio Z/K. Similar to the age-based catch curve, the length based catch curve can be used to estimate the ratio Z/K. Since in contrast to age, growth is not constant, it takes a fish longer to grow through the larger length classes. Thus, the slope of the log-converted length frequency distribution is curved and a function of time, mortality and the changing value of fish length. For the von Bertallanfy growth model, the slope of the length frequency distribution, i.e. the relationship between numbers of individuals by size class versus size class, can be estimated as;

$$\log(N_t) = \alpha + (Z/K)\log(L_\infty - l) \ldots \qquad (6)$$

Where N_t is the numbers at length *l*, and *K* and L_∞ are parameters of the von Bertalanffy growth equation.

Beverton and Holt (1957) provided the classical estimator by:

$$Z/K = L_\infty - L_{mean} / L_{mean} - L_c \quad \ldots \qquad (7)$$

where only the mean length L_{mean} in a sample of fish above the length at first capture L_c and an estimate of L_∞ are required.

This model assumes that the length frequency distribution is not truncated, i.e. the maximum age in the catch t_λ is the same as the theoretical maximum age t_{max}. This equation overestimates *Z* when the maximum age in the catch is less than that in the population (Ehrhardt and Ault 1992). This bias decreased with higher fishing mortalities due to the truncation of the length frequency distribution. This bias was corrected by calculating automatically the mean lengths for lengths greater or equal to L_c in order to truncate the length frequency distribution (Pauley and Arreguin-Sanchez 1995) and estimates of standard error were made by bootstrapping length data greater or equal to L_c.

Data Entry and Analysis

The data was entered into a Microsoft Access database specifically set up for the port sampling program and analysed using the statistical software package R (R Core Development Team 2012) and Microsoft Excel ©.

Data on date of sampling, vessel name, position of storage well, net number, well layer, and fork length (cm) were some of the fields collected. This preliminary analysis focused on the data for skipjack (SKJ) and yellowfin (YFT) and bigeye (BET) data only, by-catch data was not included.

Results

Log transformed length frequencies were plotted against length intervals to ascertain the general shape of the length distribution (Figure 1). The left side of the distribution is assumed to be truncated to the size at first capture thus the underlying distribution of length frequency was maintained as being normal.

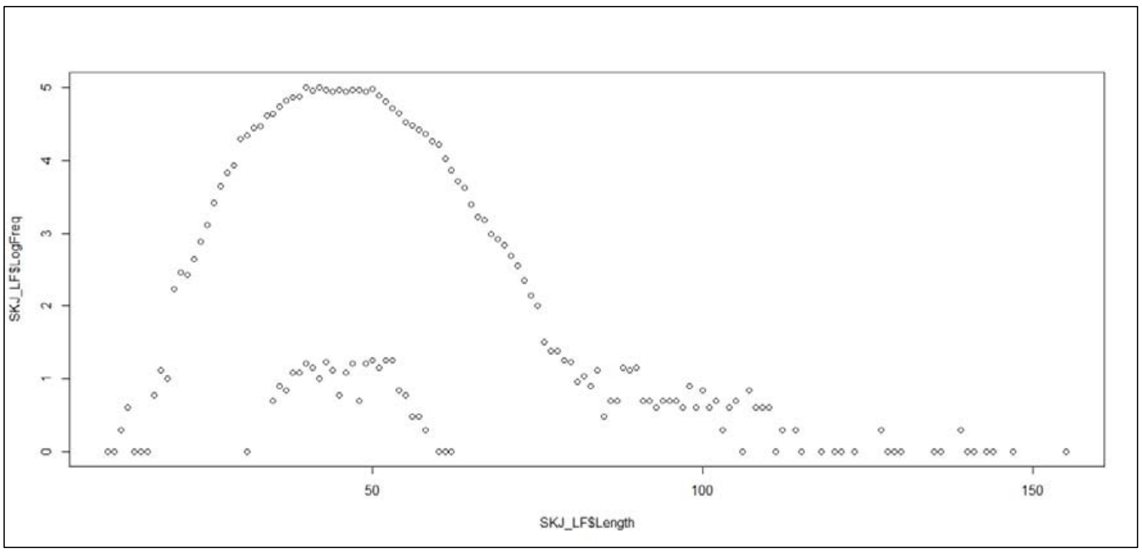

Figure 1. Scatter plot of log transformed length frequencies for skipjack tuna (*Katsuwonus pelamis*) from port sampling in Papua New Guinea.

Using the method by Shepherd (1987), estimates for t_0 were calculated for the length distribution by equating C to zero for non-seasonal growth over a twelve month period (Table1, Table 2). S_{max} represents the goodness-of-fit for the L_{inf} and K values (Table1), for which the maximum value indicates the best t_0 value to select from Table 2.

L_{inf} and K being selected in this case as 150cm and 0.3 respectively, and t_0 selected as 0.14. Length was calculated for daily length interval and mid-length interval bins for up to five thousand days using equation (3). The length data was then converted to the corresponding age data using equation (4). Table 2 shows the dissection of length frequency data to numbers-at-age, using the Shepherd (1987) method over a twelve month period.

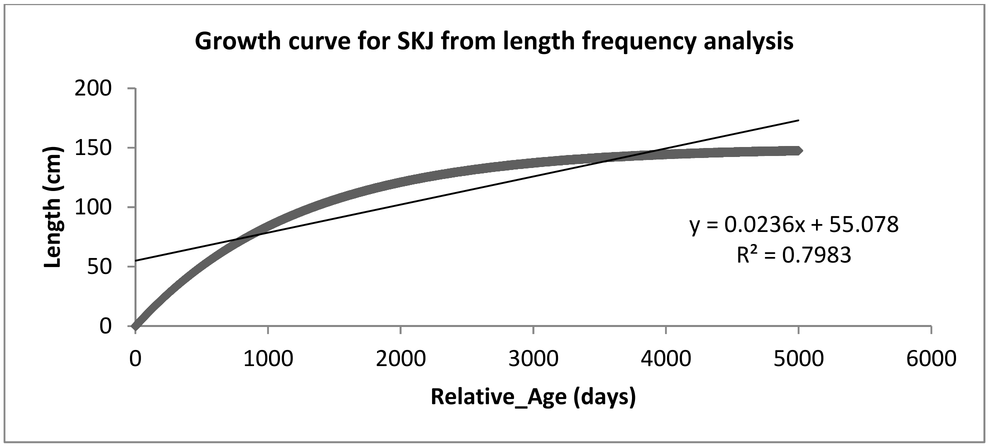

Figure 2. Growth curve for skipjack tuna (*Katsuwonus pelamis*) from length frequency analysis of port sampling data.

An age based catch curve was constructed using equation (5) and estimates of mortality calculated using equation (7).

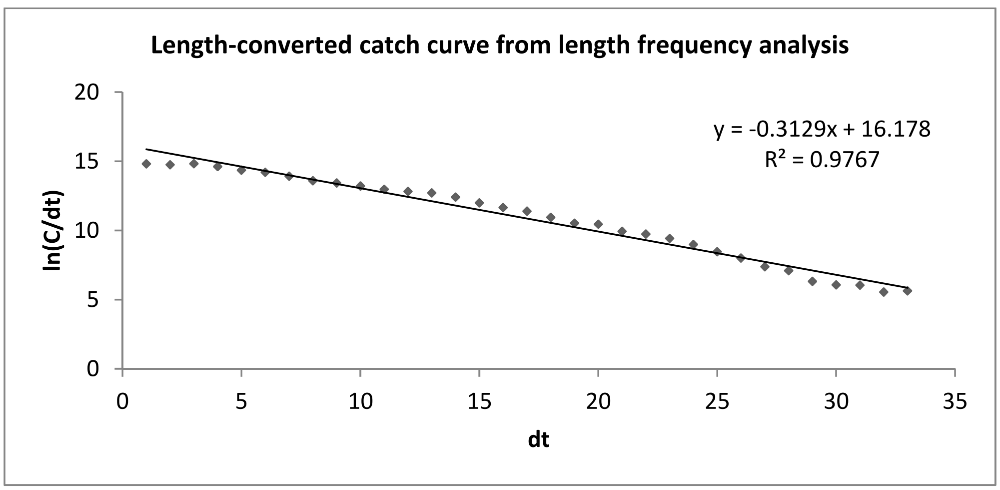

Figure 3. Length-converted catch curve for skipjack tuna (*Katsuwonus pelamis*) from length frequency analysis of port sampling data.

Discussion

From the results above the growth parameters for skipjack tuna (*Katsuwonus pelamis*) was estimated as L_{inf}=150cm, K=0.3 and t_0=0.14. The fishing mortality was estimated at F=-0.31. The fit of the length converted catch curve was R^2=0.97, while for the growth curve R^2=0.79. Year to year variability in recruitment and growth parameters will influence the results obtained in the first part of this analysis, which is highly dependent on the extent to which modes associated with individual cohorts in these frequencies are distinguishable (Majkowski, et al. 1987). From Table 2 it is evident that the majority of the skipjack catch is of the second and third age-classes. Interestingly though there is a spattering of larger age classes throughout the 12 month sapling period.

A small age sampling program is currently underway to verify the growth curve and hence develop an age converted catch curve. The next stage is to perform Virtual Population Analysis (VPA) or length-based cohort analysis, which requires more information in order to provide the basis for the formulation of fisheries management strategies. This work can be considered as the basis of providing an initial look at developing a more formal in-country analysis framework of PNG tuna catches and effort.

Acknowledgements

We would like to acknowledge the work of our port samplers and the port sampling coordinators who toil very hard to collect this data for us. We would also like to make special acknowledgement of the Japanese Trust Fund for supporting our port sampling program in 2011 and now that support has been extended to our tuna growth and ageing study.

Table 1. Values of *Smax* from range of Linf and K values using Shepherds (1987) method.

Linf\K	**0.1**	**0.12**	**0.14**	**0.16**	**0.18**	**0.2**	**0.22**	**0.24**	**0.26**	**0.28**	**0.3**
90	0.21	0.38	0.85	0.54	0.05	0.42	0.69	0.77	0.73	0.89	0.03
100	26.7	39.0	139.2	247.8	172.9	30.4	241.0	623.2	1060.3	1503.0	1881.6
110	41.7	146.0	237.6	141.2	125.0	585.3	1104.1	1608.3	2010.9	2306.5	2517.0
120	117.5	224.0	153.9	251.0	820.5	1411.0	1899.5	2249.3	2491.4	2652.7	2734.5
130	204.4	211.1	240.6	892.8	1538.4	2918.9	2360.8	2574.8	2691.3	2751.8	2787.0
140	228.1	163.0	800.7	1507.4	2039.0	2392.4	2597.3	2692.3	2756.5	2810.8	2852.8
150	222.0	542.3	1345.3	1963.0	2370.5	2579.6	2687.9	2767.1	2822.4	2873.1	2923.8

Table 2. Values of \`t_0 from range of Linf and K using Shepherds (1987) method.

Linf\K	**0.1**	**0.12**	**0.14**	**0.16**	**0.18**	**0.2**	**0.22**	**0.24**	**0.26**	**0.28**	**0.3**
90	0.21	0.38	0.85	0.54	0.05	0.42	0.69	0.77	0.73	0.89	0.03
100	0.25	0.75	0.63	0.17	0.55	0.71	0.61	0.83	0.93	0.48	0.61
110	0.48	0.52	0.16	0.58	0.34	0.67	0.94	0.15	0.33	0.48	0.61
120	0.17	0.01	0.53	0.31	0.70	0.99	0.22	0.4	0.56	0.69	0.80
130	0.70	0.35	0.19	0.66	0.98	0.23	0.43	0.60	0.73	0.84	0.94
140	0.03	0.88	0.53	0.92	0.21	0.43	0.61	0.75	0.86	0.96	0.05
150	0.37	0.30	0.79	0.13	0.39	0.59	0.74	0.87	0.97	0.06	0.14

Table 3. Numbers-at-age by dissection of length classes into age-lasses by the Shepherd (1987) method.

	Age Class / Months											
	1	2	3	4	5	6	7	8	9	10	11	12
0	2.0	2.0	0.8	0.0	0.0	15.6	7.8	44.7	24.7	115.9	1451.5	1363.0
1	1170.1	1830.5	5469.4	5627.2	4302.9	14022.2	23960.1	20644.8	16767.2	30479.3	47517.1	82009.2
2	119566.4	38945.3	104701.4	178076.7	60524.1	119491.5	102365.4	58710.6	81222.5	64910.0	44154.7	18745.2
3	37266.5	10005.0	39879.5	37548.7	16441.9	35290.6	15830.4	9200.5	14760.9	9441.6	4129.2	2156.8
4	4525.1	511.2	851.3	1703.6	466.5	904.2	342.6	130.4	324.7	252.6	513.9	238.7
5	116.0	2.0	9.6	19.4	8.1	9.8	6.7	4.1	10.6	7.5	7.7	3.2
6	1.8	0	5.0	3.5	2.5	3.0	8.1	4.1	2.4	4.0	0	0
7	2.0	0	0.8	2.0	0	3.0	1.9	1.8	2.3	0	0	0
8	1.0	0	0.2	0	1.0	0	0	0.3	0.7	1.0	0	0
9	0	1.0	2.0	0	0	0.3	7.1	1.7	0	1.0	0	0
10	0	0	0	0	0	1.7	3.5	0	0	0	0	0
11	0	0	0	0	0.6	0	3.4	0	0	0	0	0
12	0	0	1.0	0	0.4	0	4.0	0	0	0	0	0
13	0	0	0	0	0	0	2.0	0	0	0	0	0
13+	0	0	0	0	0	0	1.0	0	1.0	0	0	0

References

Ehrhardt, N. M, and J. S Ault. "Analysis of two length based mortality models applied to bounded catch length frequencies." *Transactions of the American Fisheries Society 121*, 1992: 115-122.

Hasselblad, V. "Estimation of parmmeters for a mixture of normal distributions." *Technometrics 8*, 1966: 431-444.

Kumasi, B, T Usu, L Baje, and L Kumoru. *Preliminary analysis of length frequency for FAD associated catch in the archipelagic waters of Papua New Guinea from port sampling data.* Nuku'alofa: WCPFC-SC6-2010/FT-IP-06, 2010.

Kumoru, L. *Annual report to the Commission; Information on fisheries, research and statistics.WCPFC-SC5-AR/CCM-18.* Port Moresby: Western and Central Pacific Fisheries Commission Scientific Committeee Meeting, 2008.

Majkowski, J, J Hampton, R Jones, A Laurec, and A. A Rosenberg. "Sensitivity of Length-Based Methods for Stock Assessment: Report of Working Group III." *In D. Pauly and G. R. Morgan (eds) Length-based methods in fisheries research. ICLARM Conference Proceedings 13, 468p.* International Center for Living Aquatic Resources Management, Manila, Phillipines, and Kuwait Institute for Scientific Research, Safat, Kuwait, 1987. 363-372.

Pauley, D, and F Arreguin-Sanchez. "Improving Shepherd's length composition analysis (SLCA) method for growth paraeter estimations." *Naga, the ICLARM Quarterly 18(4)*, 1995: 31-33.

Quinn, Terence. J, and Richard. B Deriso. *Quantitative Fish Dynamics.* New York: Oxford University Press, 1999.

R Core Development Team. "R: A language and environment for statistical computing." *R Foundation for Statistical Computing, Vienna, Austria* (R Foundation for Statistical Computing), 2012.

Shepherd, J. G. "A weakly parametric Method for Estimating Growth Parameters from Length Composition Data." *Poceedings of the International Conference on the Theory and Application of Length-Based Methods for stock Assessment.* Vallo: International Center for Living Aquatic Resources Management, Manila, Phillipines, and Kuwait Institue for Scientific Research, Safat, Kuwait, 1987. 113-120.

Sparre, P., and S. C Venema. *Introduction to Tropical Fish Stock Assessment, Part 1-manual.* Fisheries Technical Paper306-1. 2. 385p, Manila: FAO Fisheries Technical Paper306-1rev 2, 1992.

Ziegler, P.E, D.C Welsford, and A.J Constable. "Length-Based Assessments Revisited - Why Stock Status and Fishing Mortality of Long Lived Species Such As Toothfish Cannot Be Inferred From Length Frequency Data Alone." *CCALMR Science, Vol. 18*, 2011: 57-73.