

Estimation of Bottomfish CPUE using the Delta Method and HDAR Logbooks 1948-2010.

Kevin Piner and Hui-Hua Lee
NOAA Fisheries
PIFSC
2570 Dole Street
Honolulu, HI 96822

Abstract

An index of Hawaiian Bottomfish CPUE was created from the Hawaiian commercial logbooks. The data available included observations of catch from 4 islands, 12 months and spans the years 1948-2010. Although there were relatively few zero catch records, the index was created using the delta approach where a separate model is used to estimate proportion of positive catches and another the positive catch rate. All main effects tested were significant and included in the final model. A significant Year*Island interaction was detected that could indicate some island specific dynamics, but plots of island-specific trends did not strongly support this hypothesis. A single model estimating CPUE including all 4 islands was constructed. Estimates of year-specific variability were derived from a jackknifing procedure. Uncertainty was estimated to be very low and this is likely the result of the very large sample size.

Introduction

Analysis of fishery catch per unit effort (CPUE) often centers around the question of how to treat non-positive catch observations. Traditional error distribution assumptions do not adequately deal with typical fishery data where often 10% or more of the observations have no catch. Often a small constant is added to all catches to eliminate the zero in the numerator, but the effect of the constant on the final trend in CPUE is often poorly understood. The delta method (Lo et al. 1992) accounts for the zero catches, without relying on the addition of a constant to the catch rate, by modeling the proportion of positive tows with a binomial error distribution and the estimate of positive catch rate is derived from a separate model that uses an appropriate error distribution (Dick 2004).

The objectives of this paper were to generate an index of abundance from Hawaiian logbook data specifically using the delta approach. The request is part of a larger effort undertaken to assess the population status of bottomfish in Hawaii, and as part of that effort the methods used (Delta) were determined in advance.

Materials and Methods

Data from the HDAR logbooks of deep handline gear s were compiled and pre-processed by PIFSC staff to be consistent with other analysis being performed (pers. comm. Jon Brodziak). Available data fields included catch per trip (lbs), fishing area, island, month,

license and year. A total of 90 levels of area were present in the data. Area was defined as HDAR defined zones. There were 4 levels of island groups including: island=1 is Hawaii; island=3 is Maui, Molokai, Lanai; island=4 is Oahu; island=5 is Kauai. Hawaii had 62,000, Maui 68,000, Oahu 21,000 and Kauai 15,000 observations. Observations were spread fairly evenly through all 12 months with each winter month recording >10% of total observations and summer months roughly 5% of the total. A total of 5250 licenses were present in the data. Prior to 1993 license number changed and after 1993 license number was constant for vessels. After 1993 a total 2199 licenses appear in the records. A total 64 years were available for analysis, and years were defined as “fishing year”. Fishing year begins July 1st of the prior year and ends June 30 of the fishing year. A total of 166,992 observations were available from 1948-2011. CPUE was defined as the catch per trip.

To calculate the proportion of positive records we used a model assuming a binomial error distribution:

$$Y_{ijk} = \text{Overall mean} + \text{Year}_i + \text{Month}_j + \text{Island}_k + \text{error}$$

Where $i=1948-2011$, and $j=1-12$ months and $k=1-4$ islands.

The catch rates of the positive tows were modeled assuming a lognormal error distribution:

$$\text{Ln}(\text{CPUE})_{ijk} = \text{Overall mean} + \text{Year}_i + \text{Month}_j + \text{Island}_k + \text{error}$$

Where $i=1948-2011$, and $j=1-12$ and $k=1-4$.

The final estimate of the annual abundance index was the product of the marginal year means with appropriate bias correction (Lo et al. 1992). The variance estimates were obtained by jackknifing the data (pers. comm. E.J. Dick.).

Results and Discussion

A majority of the observations (98%) were positive catches of the deep 7 species, with a slight decline in percent positive after 1990 (Figure 1). Although the purpose of this work was to estimate CPUE trends using a Delta method, the authors note some Hessian instability of the Binomial mode used to estimate the proportion positive trips. This is likely the result of most years containing all positive observations. It is quite likely that a simpler model that adds a constant to the zero catch records would likely produce similar results without the added complexity of the delta method.

Because of the changing meaning and large number of Licenses, it was not considered a potential factor in the analysis. Because of the large number of levels of HDAR Areas (90 levels) and Years (61 levels); introducing spatial effects using Area would have reduced the number of observations to a rough average of 30 per Area / Year combination (less for Oahu and Kauai). Instead the authors choose to use Island (4 levels) as the spatial factor (instead of Area), which results in roughly 600 observations per Island and Year

(less for Oahu and Kauai) combination. Using Island along with month (12 levels) as resulted in roughly an average of 50 observations per Year, Island and Month combination. This number of observations should be enough to estimate a robust mean per statistical cell. It is noted that if population dynamics occur at smaller spatial scales we may have missed those effects in this treatment. However, it is more likely that introducing the HDAR Area into the analysis would in effect add noise rather than signal due to under-sampling at each statistical cell. Although there was not time in this particular analysis, future work at resolving an optimal (combination of HDAR Areas) may be of value (example group HDAR areas into windward and leeward sides of islands resulting in 8 levels).

The estimation of interactions in these kinds of analyses can be problematic because the large number of total observations can detect significance that does not correspond to biological reality. Furthermore, introducing multiple interactions can dramatically increase the sheer number of statistical cells thus reducing the observations per cell to very low numbers. We felt that deriving a CPUE series that is intended to represent the change in relative abundance of the population should be hypothesis driven rather than enumeration of all possible combinations. Our biological assumptions in this analysis are that bottomfish are relatively sedentary (Anon 2010) as adults and unlikely to move between islands. Thus typical interactions, such as Island*Month which might be important for a migratory species showing seasonal movements were not considered (again to keep sample size per cell relatively high). However, the Year*Island interaction was investigated as local depletion is a potential in reef populations. A real interaction of Year and Island might indicate that population trends by island were different (Maunder and Punt 2004). It was determined that the interaction was significant ($p < 0.001$) but a plot of CPUE by island did not long-term differences in trends (Figure 2a). The exception is potentially island Island 3 where a more significant decline may be identified since 1960 although all 4 islands show decline (Figure 2b). However, all islands indicated both long-term and more contemporary declining trends and thus a single index was created. The authors recommend future research in this area. Research could be directed toward evaluating the need for weighting each islands CPUE by habitat size or conducting individual stock assessments for each island.

The ANOVA tables for the models are given in Table 1. Because of the large sample sizes all effects were significant. Year-specific CPUE and associated CV are given in Table 2 and Figure 3. The large sample size resulted in estimated CV under 5% which are probably unrealistically low. Additional thought will need to be given to subjectively set the level of uncertainty in the series. Diagnostics do not indicate severe departure from model assumptions (Figures 4 and 5).

Literature Cited

- Anon. (2010) Fishing for Science. Lawaia. Issue 5 p 43-47.
Dick, E.J. (2004) Beyond 'lognormal versus gamma': discrimination among error distributions for generalized linear models. Fish. Res. 70(2-3): 351-366.

Lo, N. C. H., L. D. Jacobson, J. L. Squire. 1992. Indices of relative abundance from fish spotter data based on delta-lognormal models. *Can. J. fish. Aquat. Sci.* 49(12):2515-2526.

Maunder, M.N., and A.E. Punt. (2004). Standardizing catch and effort data: a review of recent approaches. *Fish Res.* 70(2-3):141-159

Table 1. ANOVA Table for final GLM results.

Binomial Model

Source	Deviance	DF	Square	Pr > ChiSq
Intercept	29941.6686			
year	25708.6852	63	4232.98	<.0001
month	25645.3977	11	63.29	<.0001
island	25437.6532	3	207.74	<.0001

Lognormal Positive Catch rate

Source	Likelihood	DF	Square	Pr > ChiSq
Intercept	-546015.92			
year	-541714.15	63	4301.78	<.0001
month	-540712.68	11	1001.46	<.0001
island	-521946.14	3	18766.5	<.0001

Table 2. Estimated CPUE and CV.

Year	CPUE (lbs/day)	CV	Year	CPUE (lbs/day)	CV
1948	104.9	0.03	1980	78.6	0.03
1949	73.2	0.03	1981	76.4	0.02
1950	79.6	0.03	1982	69.2	0.02
1951	93.3	0.03	1983	74.0	0.02
1952	106.0	0.04	1984	62.0	0.02
1953	97.2	0.04	1985	71.8	0.02
1954	132.4	0.04	1986	75.6	0.02
1955	194.8	0.05	1987	89.8	0.02
1956	129.8	0.05	1988	95.3	0.02
1957	161.3	0.05	1989	89.2	0.02
1958	100.0	0.05	1990	79.7	0.02
1959	66.3	0.04	1991	68.5	0.02
1960	100.6	0.04	1992	65.4	0.02
1961	157.4	0.06	1993	59.5	0.02
1962	177.1	0.05	1994	68.1	0.02
1963	107.7	0.05	1995	70.7	0.02
1964	107.2	0.04	1996	63.9	0.02
1965	129.2	0.04	1997	65.2	0.02
1966	116.1	0.04	1998	59.3	0.02
1967	110.7	0.03	1999	64.4	0.02
1968	103.8	0.04	2000	73.5	0.02
1969	99.1	0.04	2001	68.2	0.02
1970	87.0	0.04	2002	66.5	0.02
1971	79.1	0.03	2003	70.2	0.02
1972	89.2	0.03	2004	67.7	0.02
1973	87.5	0.03	2005	74.1	0.02
1974	85.3	0.03	2006	69.1	0.02
1975	84.6	0.03	2007	72.5	0.02
1976	94.2	0.03	2008	82.7	0.02
1977	79.8	0.03	2009	76.2	0.02
1978	83.1	0.03	2010	61.5	0.02
1979	90.5	0.03	2011	51.4	0.02

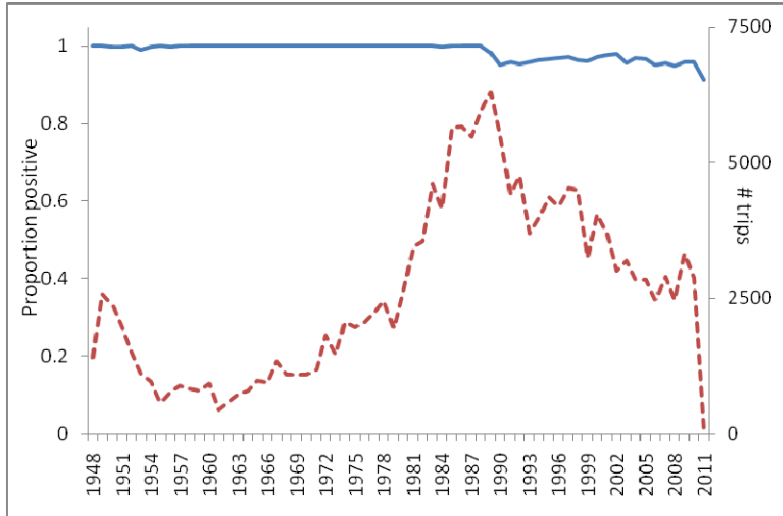


Figure 1. Number of trips (dotted line) and proportion positive (solid line) by year. Low number of trips in 2011 reflect only partial year data.

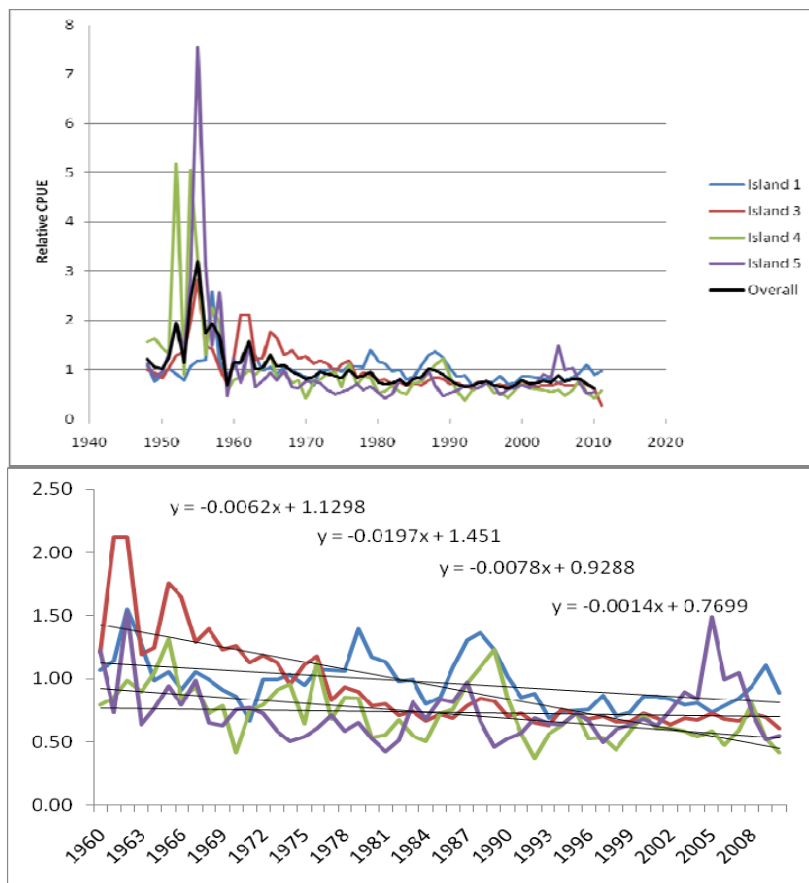


Figure 2. Estimated CPUE trends by island for the periods a) 1948-2011 and b) 1960-2011. Linear models (equation given above) are fitted to each of the areas to indicate more recent trends.

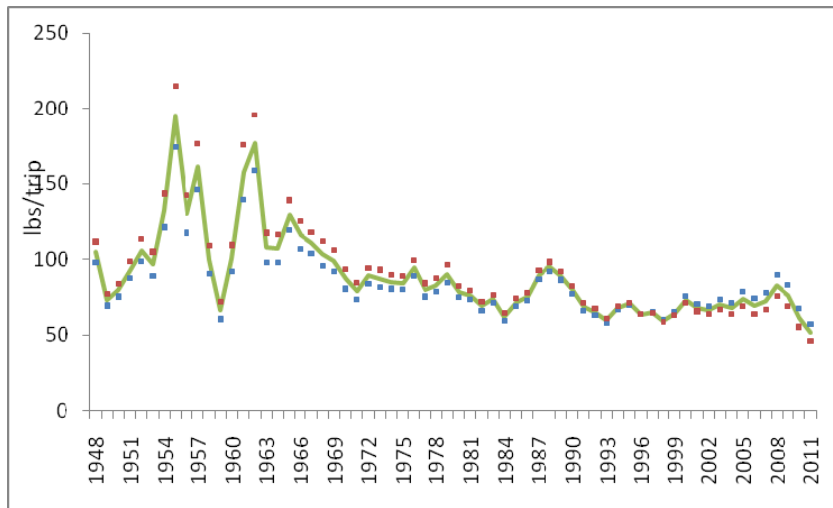


Figure 3. Estimated CPUE for all islands and years (solid line) and associated 95% CI (squares) based on the jackknifing procedure.

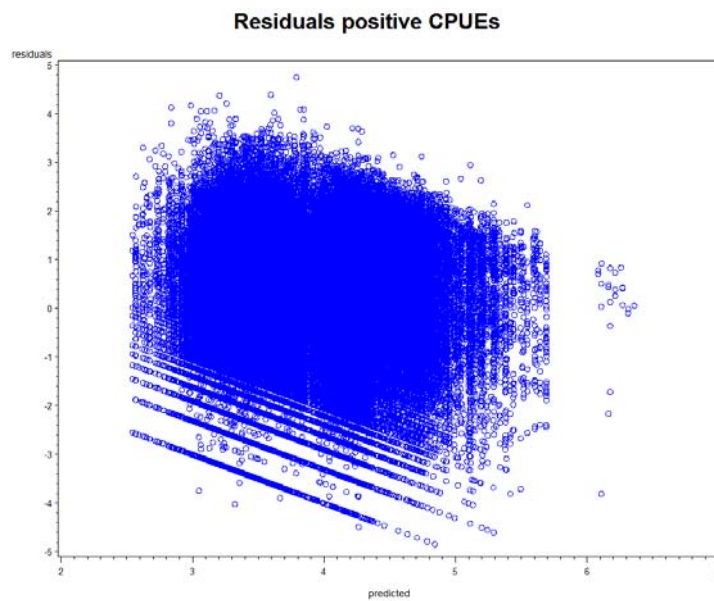


Figure 4. Plot of the predicted CPUE and residual from positive catch rate model.

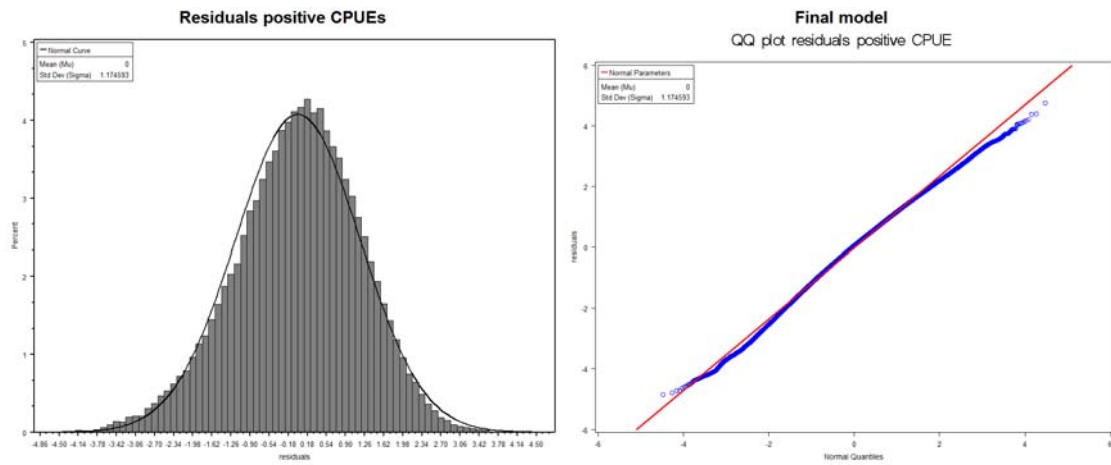


Figure 5. Histogram of a) residuals and b) QQ plot from the positive catch rate model.