Standardization of the Striped Marlin (*Kajikia audax*) Catch per Unit Effort Data Caught by the Hawaii-based Longline Fishery from 1994-2017 Using Generalized Linear Models

Sculley, Michelle[a]

a Joint Institute for Marine and Atmospheric Research, University of Hawaii
c/o National Marine Fisheries Service
1845 Wasp Boulevard
Honolulu, HI 96818

Working document submitted to the ISC Billfish Working Group, International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean, 14-21 January 2019, Honolulu, HI, USA. Document not to be cited with author's permission.

Abstract

This working paper provides the standardization of the Hawaii-based longline fishery striped marlin (*Kajikia audax*) catch per unit effort (CPUE) data. Three different distributions with up to 14 different explanatory variables were explored for the combined dataset. The delta-lognormal generalized linear mixed model (DL-GLMM) provided the best fit to the data based upon percent deviance explained. Using this best-fit model, standardizations with the shallow-set and deep-set sectors were conducted separately and compared. Results showed that the deep-set sector standardized CPUE was very similar to the combined dataset. The shallow-set CPUE series was higher than the other CPUE time series and was flat and highly variable making it a poor candidate for inclusion in the assessment model. The diagnostics of the combined dataset do not suggest any problems with poorly fitted data; therefore, it was recommended to use the combined dataset DL-GLMM standardized CPUE for the 2019 striped marlin base-case assessment model.

Introduction

Striped marlin (*Kajikia audax*) is a tropical and subtropical species of billfish found in the Pacific Ocean. It is primarily caught as a non-target species in longline fisheries targeting tuna and swordfish, although it is occasionally targeted in commercial and recreational fisheries. The most recent stock assessment of striped marlin was in 2015 of the Western and Central North Pacific stock, which was found to be overfished ($SSB < SSB_{MSY}$) and overfishing was occurring ($F > F_{MSY}$, ISC BILLWG, 2015). This was an update assessment based upon the 2011 assessment model. The billfish working group (BILLWG) of the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean (ISC) has agreed to do a benchmark assessment of the Western and Central North Pacific striped marlin in 2019. This assessment will be bounded by the boundaries of the Western and Central Pacific Fisheries Commission at 150°W latitude and the equator, which is different from the 140°W boundary from the 2015 assessment.

The Hawaii-based longline fishery catches striped marlin as non-target species in both the swordfish-targeting shallow-set sector and the tuna-targeting deep-set sector. This fishery spans from approximately 180° to 120°W, therefore only a portion of the fishery will be included in the 2019 assessment. However, approximately 90% of the catches of striped marlin occur west of 150°W, therefore the majority of the Hawaii-based fishery will be included in this assessment. This working paper details the methods and results of the standardization of striped marlin from the Hawaii-based fishery and provides additional information on the misidentification rate and discard rate of striped marlin in this fishery.

Methods

Data

The Pacific Islands Regional Fishery Observer Program (PIROP) provides detailed set-by-set data on the Hawaii-based longline fishery including catch in numbers of fish and a variety of operational variables, among them: location as latitude and longitude, vessel ID, hooks per float, total number of hooks set, type of bait used, and time longlines were set, following the procedures outline in the PIROP observer manual (Pacific Islands Regional Office, 2017). The standardization uses this data set instead of the commercial logbook data to ensure the analyses were conducted according to ISC standards on using the best available science. There is a known issue of misidentification and discards of striped marlin from the Hawaii longline fishery from commercial logbook data (Walsh *et al.,* 2007). Discards of striped marlin prior to 2004 were estimated to be 7% and striped marlin catches were estimated to be underreported by approximately 5% due to misidentification (Walsh *et al.,* 2007). It should be noted, however, that estimated misidentification rates for striped marlin varied inversely with observer coverage rates, which suggests that misidentification rates were lower after 2004 when observer rates were much higher.

Data were extracted from the PIROP database on 30 August 2018 for this analysis. There were 232,776 total records, 40% of these were positive catches. There were 284,241 fish recorded in the observer data from the deep-set sector, which accounts for 92% of the catch; 24,377 fish were recorded in the shallow-set sector or 8% of the catch.

Observers were first placed onboard Hawaii-based longline vessels in 1994. Observer coverage varied significantly prior to 2000, with observer coverage between 3.3 and 10.4% annually for the entire fishery (NMFS, 2017). Due to interactions with protected species the shallow-set sector was closed from 2001–2004. When it was reopened, 100% observer coverage was implemented on shallow-set trips and ~20% observer coverage was implemented on deep-set trips (Gilman *et al.,* 2007). The deep-set trips are typically further south than the shallow-set trips, which are concentrated around the sub-tropical frontal zone (STFZ) where large swordfish are caught (Sculley et al., 2018). After the closure, shallow sets were defined as sets with fewer than 15 hooks per float, however, prior to the closure most sets targeting tuna used 10 or more hooks per set (Figure 1). Previously, the 15 hooks per float cutoff was used for all years of the time series, however, were 567 sets that used 10–14 hooks per float before 2004 which were listed in the observer database as deep-sets targeting tuna. Therefore, it was decided that using 10 hooks per float as the division between deep and shallow sets prior to 2004 and using 15 hooks per float from 2004 through the present would adequately capture the change in the definition and behavior of the fishers.

In general, the deep-set sector has a higher encounter probability while the shallow-set sector catches more striped marlin (higher positive CPUE) when encountered (Figure 2). The combined nominal CPUE mimics the deep-set sector CPUE as the majority of the data come from the deep-set sector. The nominal CPUE for the shallow-set sector is highly variable, relatively flat, and generally higher than the deep-set and combined nominal CPUEs. Overall, CPUEs are highest west of the Hawaiian Islands and decreases over time (Figure 3). Based upon previous standardizations, the combined dataset was used for the standardization, and the final delta-lognormal model was used to provide standardized CPUE series for the shallow-set and deep-set sectors, to compare the results.

Environmental variables used in the standardization were obtained from publically available data sets. Sea Surface temperatures (SST) from January 1994 to 2017 were based on monthly 0.5° resolution composites from the NOAA GOES-E/W satellite downloaded from Pacific Islands Fisheries Science Center (PIFSC) OceanWatch (2017). Both the Southern Oscillation Index (SOI) and the Pacific Decadal Oscillation Index (PDO) were monthly region wide indices

(NOAA NCDC, 2017). Mixed layer depth (MLD) was based on 0.33° × 1° monthly means of GODAS data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA[1].

CPUE Standardization

After reviewing previous standardization attempts of the Hawaii-based longline fishery striped marlin data (Langseth 2015; Walsh and Chang 2015), it was decided that three different distributions would be used to standardize the CPUE data: delta-lognormal generalized linear mixed model (GLMM), Poisson generalized linear model (GLM), and negative binomial GLM. A zero-inflated model (Poisson or negative binomial) was not included as Langseth 2015 showed that it was not an improvement over the Poisson GLM and 40% of the data were positive catches. The delta-lognormal GLMM included 14 potential explanatory variables. Year, Quarter, Hooks per float, bait type, begin, and set type were included as factors. Sea surface temperature, mixed layer depth, latitude, longitude, the Pacific Decadal Oscillation (PDO) index, the Southern Oscillation Index (SOI), and the begin set time were included as continuous variables. Vessel, based upon the permit number, was included as a random effect to account for differences in fishers behaviors. In addition, the Poisson and negative binomial models included log(HPSet) as an offset where HPSet is the number of hooks per set. Both the Poisson and negative binomial models with Vessel as a random effect failed to converge, so no random effect was included in those distributions.

Begin is a factor with four levels describing the time of day in which the set was initially deployed with 1 = midnight – 0600, 2 = 0600-1200, 3 = 1200-1800, and 4 = 1800-2400. Set type was a factor with two levels indicating if the set was shallow or deep with shallow sets identified as sets with fewer than 10 hooks per float. Bait type is a code that indicates the type of bait used when setting the hooks; these are typically some kind of baitfish such as mackerel, squid, or a combination of baits. Begin set time was the time (in hours) the set was initially deployed. In the first round of model selection, models with set type and hooks per float and begin and begin set time were compared and the models with the lowest AIC were included in future model selection steps. For all three distributions, begin and hooks per float had lower AICs than begin set time and set type and were used in subsequent model selection steps.

Explanatory variables were added using forward stepwise selection with variables being selected based upon the lowest AIC, most deviance explained, and if they were statistically significant based upon a Chi-squared likelihood ratio test. Additional variables were not included if they were not significant based upon the likelihood test (Bigelow *et al.*, 1999). Final models for each time series are presented in Table 1.

Annual mean CPUE was calculated from the final binomial and lognormal models using the estimated marginal means package in R (emmeans, Lenth *et al*., 2017; R version 3.4.0, R Core Team, 2017) which accounts for the unbalanced nature of the data and missing values, not allowing for large numbers of observations in a level of a factor to have an undue influence on the average of the values. Annual mean CPUE was then back-transformed into normal space and bias corrected. Then the means from the binomial model were multiplied with the means from the lognormal model to obtain the final standardized annual CPUE values. Standard deviations

[1] http://www.esrl.noaa.gov/psd/

were estimated in a similar manner: individual model values were back transformed into normal space then combined for each time series based upon the Goodman (1960) estimator (Lauretta *et al.*, 2016).

Results and Discussion

Discards of striped marlin estimated from the entire observer dataset were very low (<0.5%) while those estimated for the 1994-2004 time period were approximately 7%. This discrepancy is likely due to the difficultly of identifying striped marlin before they are discarded which was accounted for in the Walsh *et al.,* (2007) analysis. Therefore, while the number of fish discarded was included in the CPUE standardization, it should be noted that these are likely underestimated and therefore CPUEs may be negatively biased. However, it is reasonable to assume that the trend of the standardized CPUE data is consistent with the trend in relative abundance despite this discrepancy.

Comparing the percent deviance explained between the three models, the delta-lognormal model explained the most deviance in the data (positive catch rates = 76% deviance explained, encounter probability = 9% deviance explained) compared to the Poisson GLM (14% deviance explained) and negative binomial GLM (10% deviance explained). This is consistent with findings by Langseth (2015) who suggested that the delta-lognormal model was the best model to use to standardize striped marlin CPUE. Final model configurations for all three models are in Table 1. Generally, operational variables explained more of the variance in the data than the environmental variables (Table 2). The correlations between CPUE and the environmental variables was generally very low (Table 3**Error! Reference source not found.**). Correlations were strongest with the spatial variables latitude and longitude (Figure 4). Higher positive catches were observed in the shallow-set sector; however, the deep-set sector was more likely to have a positive encounter (Figure 5). Higher catches were also seen during dusk and dawn and for bait types 30 and 56, which correspond to opelu (mackerel scad) and opelu with sanma (Pacific saury) types of bait which suggests that striped marlin prefer mackerel scad as those are also the only two bait types which include mackerel scad. However, mackerel scad was only used in about 100 sets. Only two environmental variables suggested trends related to catch rates: there tended to be higher catch rates at SSTs warmer than 24°C and at MLDs shallower than 100m (

Figure 6). Neither PDO nor SOI showed any obvious trends with CPUE.

The standardized combined CPUE trend was very similar to the nominal CPUE trend, with the standardized values slightly lower than the nominal values (Figure 7, Table 4). Comparing the standardized CPUEs from the combined dataset and the separated shallow-set and deep-set trends, the deep-set CPUE is very similar to the combined CPUE in both scale and trend (Figure 8). However, the shallow-set standardized CPUE is very different. Values tend to be highly variable and higher than those when using the combined dataset, as well as relatively flat rather than a downward trend. This type of CPUE series is generally not useful in a stock assessment model as it provides very little information about the relative abundance of the stock due to the high annual variability, and would typically not be included in the base-case model. Therefore, it is recommended to use either the combined CPUE series or only the deep-set CPUE series for the 2019 striped marlin assessment. Furthermore, the residual plots from the positive catches lognormal GLMM does not indicate any patterns related to the type of set (Figure 9), and it could be concluded that using the combined CPUE series would be reasonable for this assessment.

Diagnostics for the delta-lognormal GLMM with the combined dataset show no significant deviations from the assumption of normality. Pearson residuals for the positive catch lognormal model appeared to be randomly distributed around zero and only deviated from the normal Q-Q line at the extremes of the dataset (Figure 9). When the residuals are compared to each explanatory variable, there appears to be a slight negative bias for the operational variables (Figure 10Figure 11) that was consistent in the deep-set only and shallow-set only CPUE standardizations. Residuals plotted against the environmental variables do not appear to show any significant patterns (Figure 12). Diagnostics for the encounter probability binomial model also show little patterns except at the extremes of the data (Figure 13). The binned residual plot shows that except at these extremes when the residuals tend to be positive, the majority of the residuals fall within the 95% confidence interval, indicating a good fit. The plots of the quantile residuals compared to each explanatory variable do not appear to be biased, but generally have a median value of zero (Figure 14Figure 15).

The CPUE standardization from the 2015 assessment (Walsh and Chang, 2015) has a very similar trend and pattern to the standardization prepared for the 2019 assessment (Figure 17). The 2015 standardization is more variable and is lower in the 2000s but overall provides the same information as the 2019 standardization.

Conclusions

While there is likely some bias in the estimates of CPUE due to the problems in misidentification and discards, these data are the best available science and are likely consistent with the trends in abundance of the striped marlin available to the Hawaii-based longline fishery. The best-fit model was the delta-lognormal generalized linear model, which explained 76% of the deviance in the positive catch rates and 9% of the deviance in the encounter rates. It is recommended to use the combined dataset for the standardized CPUE values in the stock assessment model, as the diagnostics do not show any significant patterning between the deep-set data and the shallow-set data. Furthermore, the combined dataset trends are consistent with the deep-set CPUE time series, and the shallow-set time series would not be a useful indicator of relative abundance as it is highly variable and relatively flat. It is interesting to note that the environmental variables included in this standardization do not appear to be highly correlated to striped marlin CPUE and additional research should be done to identify any environmental covariates that may be important to striped marlin catch rates.

Literature Cited

Bigelow, K. A., C. H. Boggs and He, X. I. 1999. Environmental effects on swordfish and blue shark catch rates in the US North Pacific longline fishery. Fisheries Oceanography 8(3): 178-198.

Gilman, E., D. Kobayashi, T. Swenarton, N. Brothers, P. Dalzell and Kinan-Kelly, I. 2007. Reducing sea turtle interactions in the Hawaii-based longline swordfish fishery. Biological Conservation 139(1–2): 19-28.

Goodman, L. A.1960. On the exact variance of products. Journal of the American Statistical Association. 55(292): 708- 713.

ISC BILLWG. 2015. Report of the billfish working group workshop (Annex 10). International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean 20-28 April 2015. Yokohama, Japan.

Langseth, B. (2015). Standardization of striped marlin (*Kajikia audax*) CPUE in the Hawaiian longline fishery: II. Additional covariates, distribution, and use of data from the deep-set fishery sector only. ISC/15/BILLWG-01/06.

Lauretta, M. V., Walter, J.F., and Christman, M.C. 2016. Some considerations for CPUE standardization; variance estimation and distributional considerations. Collect. Vol. Sci. Pap. ICCAT, 72(9): 2304-2312.

Lenth, R., Love, J. and Herve, M. 2017. Estimated Marginal Means, aka Least-Squares Means. R Package. Version 1.0 https://github.com/rvlenth/emmeans

NMFS. 2017. Hawaii longline fishery logbook statistics -non-confidential summary tables. Available online at http://www.pifsc.noaa.gov/fmb/reports.php, accessed 8 May 2017. National Marine Fisheries Service, Pacific Islands Fisheries Science Center, Honolulu.

NOAA NCDC 2017. Spatially and temporally large-scale anomalies that influence the variability of the atmospheric circulation. Online Database https://www.ncdc.noaa.gov/teleconnections/ Accessed 10 October 2017.

Pacific Islands Fisheries Science Center (PIFSC) OceanWatch. (2017) Central Pacific Node. http://pifsc-oceanwatch.irc.noaa.gov/ Accessed: 11 October 2017.

Pacific Islands Region Office (PIRO). 2017. Hawaii Longline Observer Program Observer Field Manual. Version LM.17.02. National Oceanic and Atmospheric Administration, Pacific Islands Region, Honolulu, Hawaiʻi.

R Core Team 2017 R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

U.S. Naval Observatory. 2017. Fraction of the Moon Illuminated at Midnight Hawaii-Aleutian Standard Time. http://aa.usno.navy.mil/data/docs/MoonFraction.php Accessed 10 October 2017.

Walsh, W. A., Bigelow, K. A., and Ito, R. Y. (2007). Corrected Catch Histories and Logbook Accuracy for Billfishes (Istiophoridae) in the Hawaii-based Longline Fishery. NOAA Technical Memorandum NMFS-PISCF-13. 49 pp.

Walsh, W.A. and Chang, Y.J. (2015). Standardization of Striped Marlin *Kajikia audax* CPUE for the Hawaii-based Longline Fishery during 1995–2013 using Generalized Linear Models: An Update from 2011. ISC/15/BILLWG-01/03.

Tables

Table 1. Final models and percent deviance explained for each distribution tested.

Model		**% Deviance Explained**
Positives DL-GLMM	Log(CPUE) ~ Year + HPF + Quarter + Bait + Lon + Lat + Begin + MLD + PDO + SOI + SST	76%
Encounter probability	Proportion Positive ~ Year + Quarter + Begin + Lon + MLD + Lat + Bait + HPF + SOI + PDO	8.6%
Poisson	Number of Fish Caught ~ Year + Quarter + HPF + Lat + Lon + MLD + Begin + Bait –offset (log(HPSet))	17.3%
Negative Binomial	Number of Fish Caught ~ Year + Quarter + Bait + MLD + HPF + Lat + Lon + Begin + PDO + SOI + SST – offset(log(HPSet))	10.5%

Table 2. Percent deviance explained by each parameter for the Delta-Lognormal GLMM model components: the lognormal model on the positive catches and the binomial model on the encounter rate.

Lognormal Model			**Binomial Model**		
Parameter	Deviance	% Deviance	Parameter	Deviance	% Deviance
Year	23663	24%	Year	344183	2.9%
Quarter	30858	0%	Quarter	348612	1.6%
HPF	24665	20%	HPF	351220	0.9%
Set Type	25431	18%	Set Type	352057	0.7%
Bait	25274	18%	Bait	351313	0.9%
Begin	26403	15%	Begin	351063	0.9%
BeginSetTime	28152	9%	BeginSetTime	351950	0.7%
SST	30975	0.03%	SST	354351	0.01%
MLD	30761	1%	MLD	354378	0.001%
Lat	30551	1%	Lat	353450	0.3%
Lon	30650	1%	Lon	348065	1.8%
PDO	30923	0.2%	PDO	353415	0.3%
SOI	30974	0.03%	SOI	354017	0.1%

Table 3. Correlations and p-values between striped marlin CPUE and candidate environmental and spatial variables.

Parameter	**MLD**	**SOI**	**PDO**	**SST**	**Begin Set Time**	**Lon**	**Lat**
CPUE	-0.0777	0.0199	0.047	-0.0203	0.283	-0.136	0.0932
p-value	<2.2e-16	1.53E-10	<2.2e-16	6.37E-11	<2.2e-16	<2.2e-16	<2.2e-16

Table 4. Nominal and standardized CPUE values and CVs for the combined Hawaii-based longline fishery.

Year	Standardized CPUE	CV	Nominal CPUE
1995	1.19	0.78	1.88
1996	0.98	0.89	1.61
1997	0.74	1.06	1.24
1998	0.83	1.00	1.30
1999	0.70	1.10	1.17
2000	0.55	1.32	0.85
2001	0.72	1.07	1.00
2002	0.42	1.60	0.56
2003	0.90	0.90	1.26
2004	0.55	1.27	0.77
2005	0.53	1.30	0.77
2006	0.54	1.28	0.75
2007	0.30	2.02	0.37
2008	0.39	1.61	0.64
2009	0.25	2.28	0.38
2010	0.18	2.91	0.20
2011	0.33	1.77	0.48
2012	0.28	2.02	0.46
2013	0.25	2.19	0.42
2014	0.35	1.71	0.55
2015	0.31	1.86	0.50
2016	0.30	1.89	0.44
2017	0.26	2.08	0.41

Figures

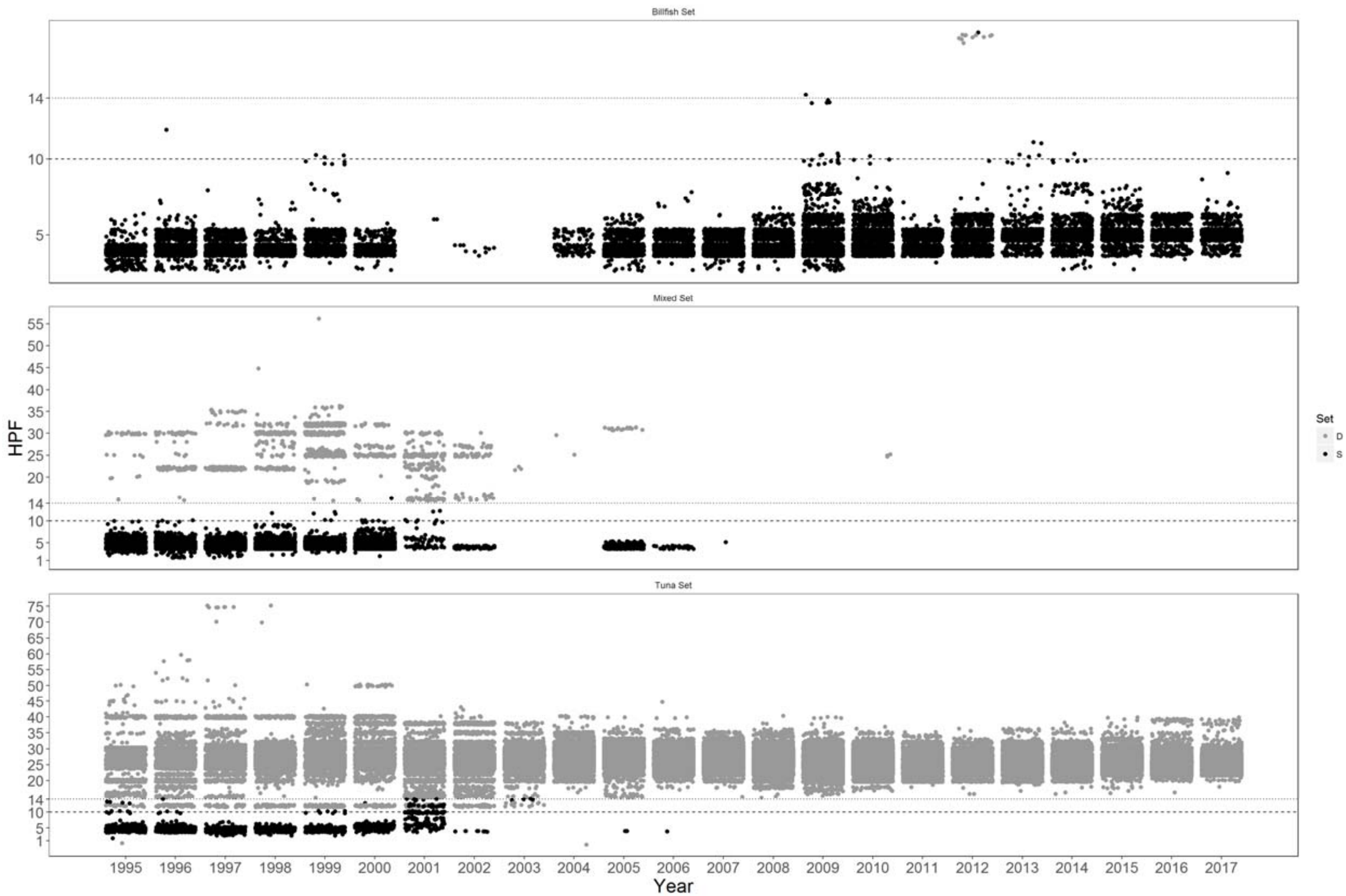

Figure 1. Hooks per float (HPF) vs Year and Target, dotted line indicates 14 HPF, dashed indicates 10 HPF, gray indicates observer recorded deep set, black indicates observer recorded shallow set.

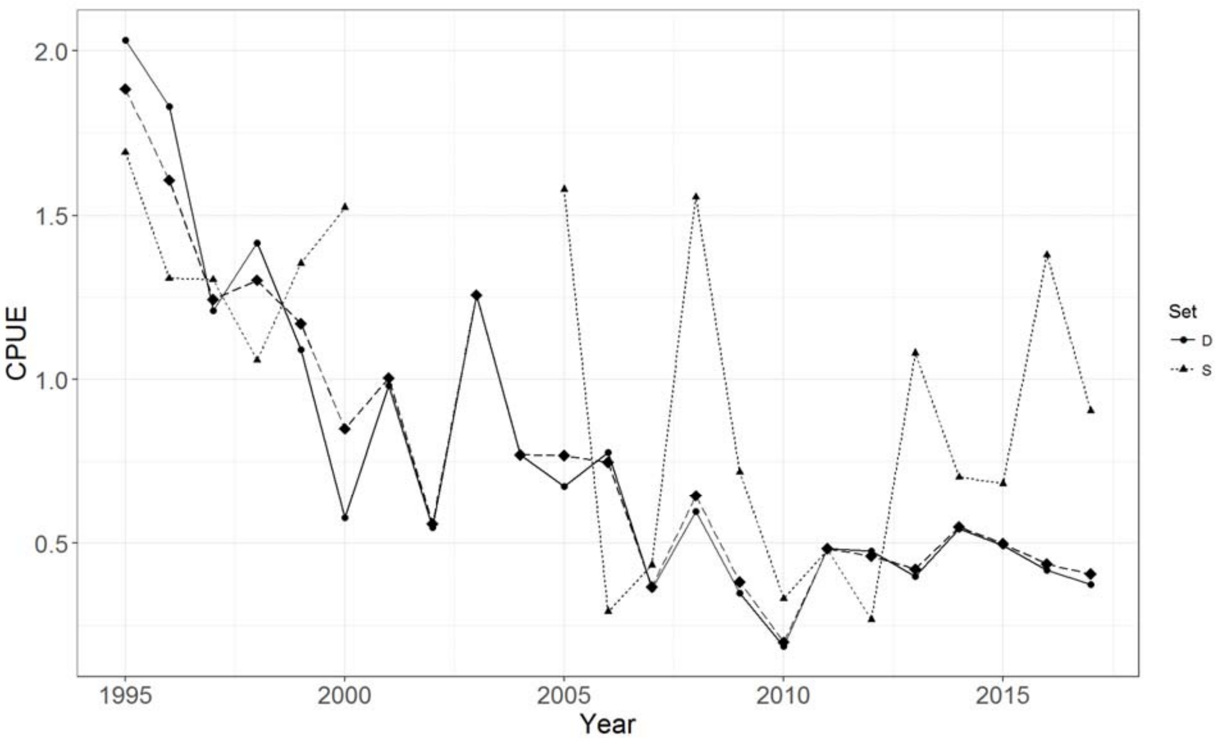

Figure 2. Nominal CPUE by set determined by a 10 HPF cutoff for shallow sets prior to 2004 and 14 HPF after 2004. Dashed diamond = combined CPUE; dotted triangles = shallow-set (S) only CPUE; solid circles = deep-set (D) only CPUE.

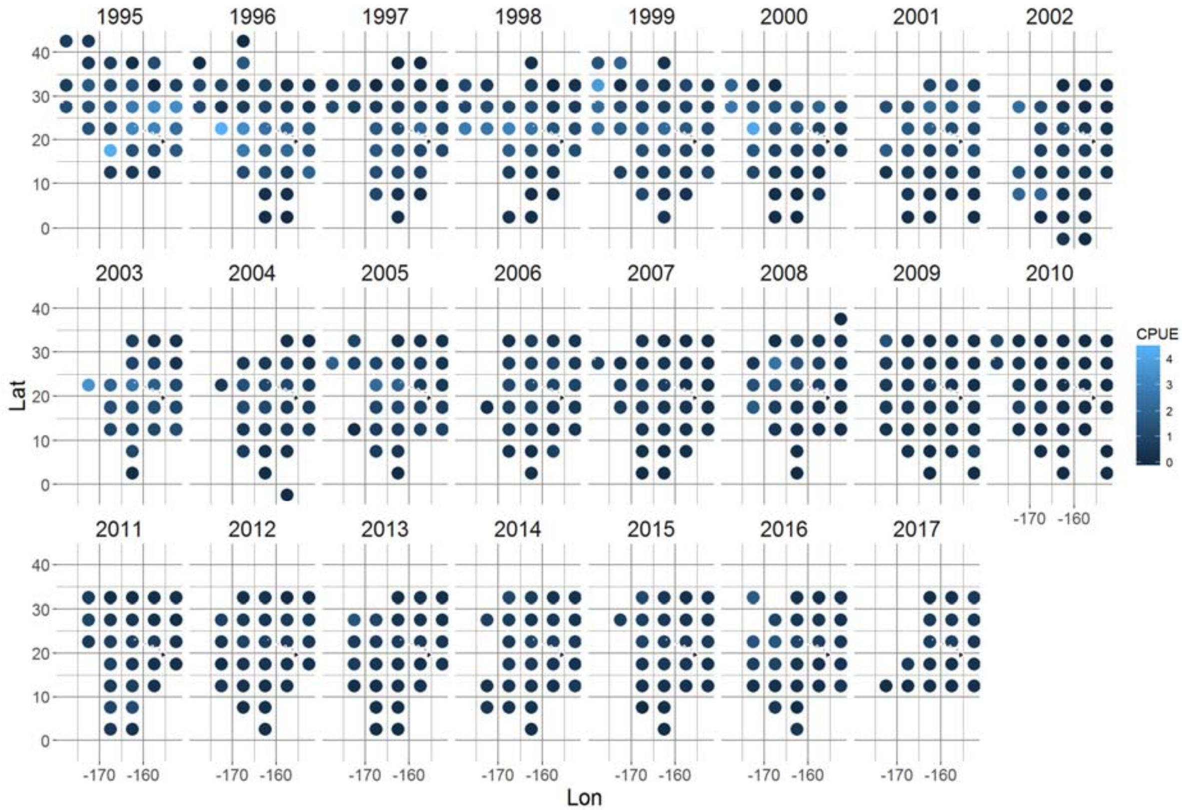

Figure 3. Nominal CPUE from the combined dataset by year and 5x5 degree location. Cells with fewer than three vessels conducting fishing were removed for confidentiality.

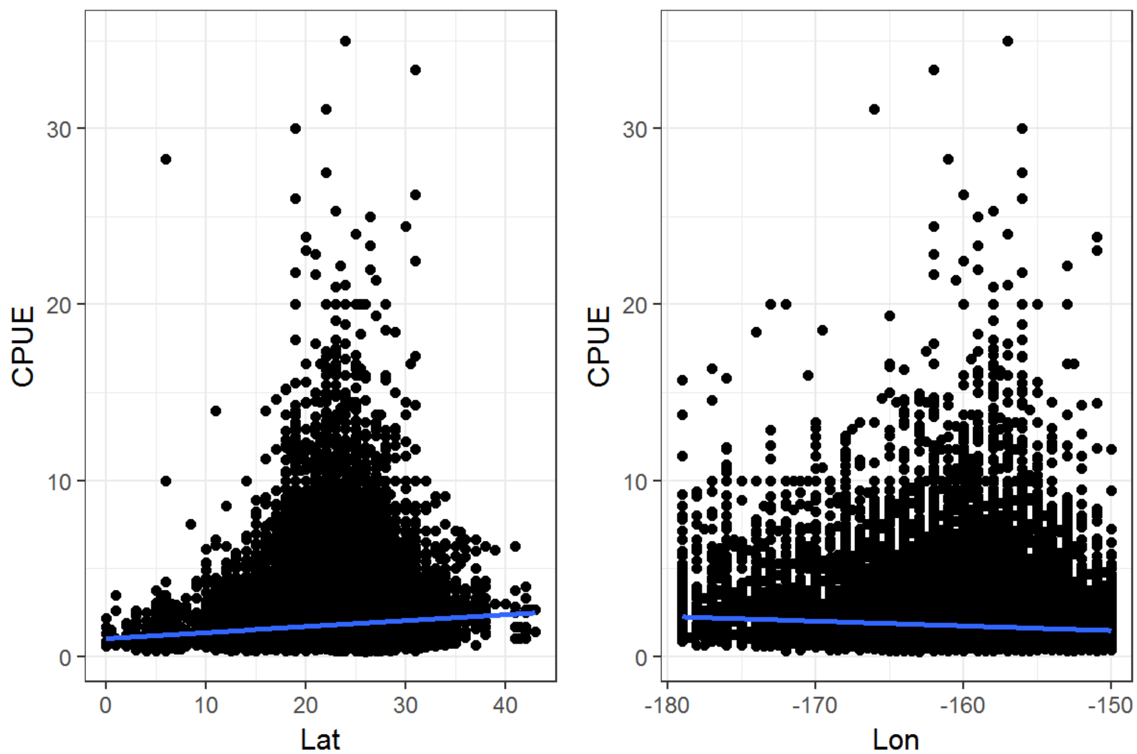

Figure 4. Nominal positive CPUE for the combined dataset vs latitude and longitude. Blue line indicates the trend of a Generalized Additive Model.

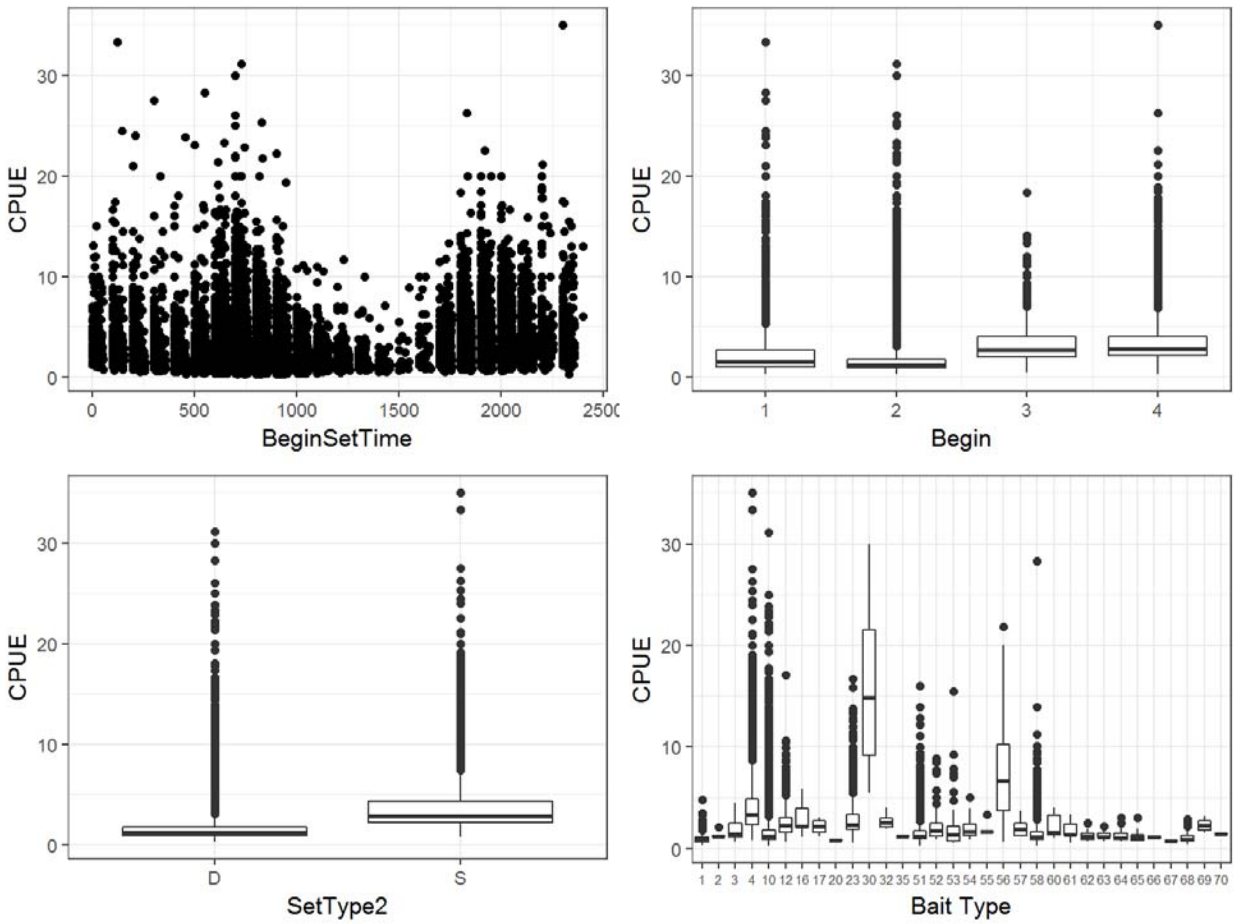

Figure 5. Nominal positive CPUE versus Begin Set Time (upper left), Begin (upper right), Set type (lower left), and Bait type (lower right).

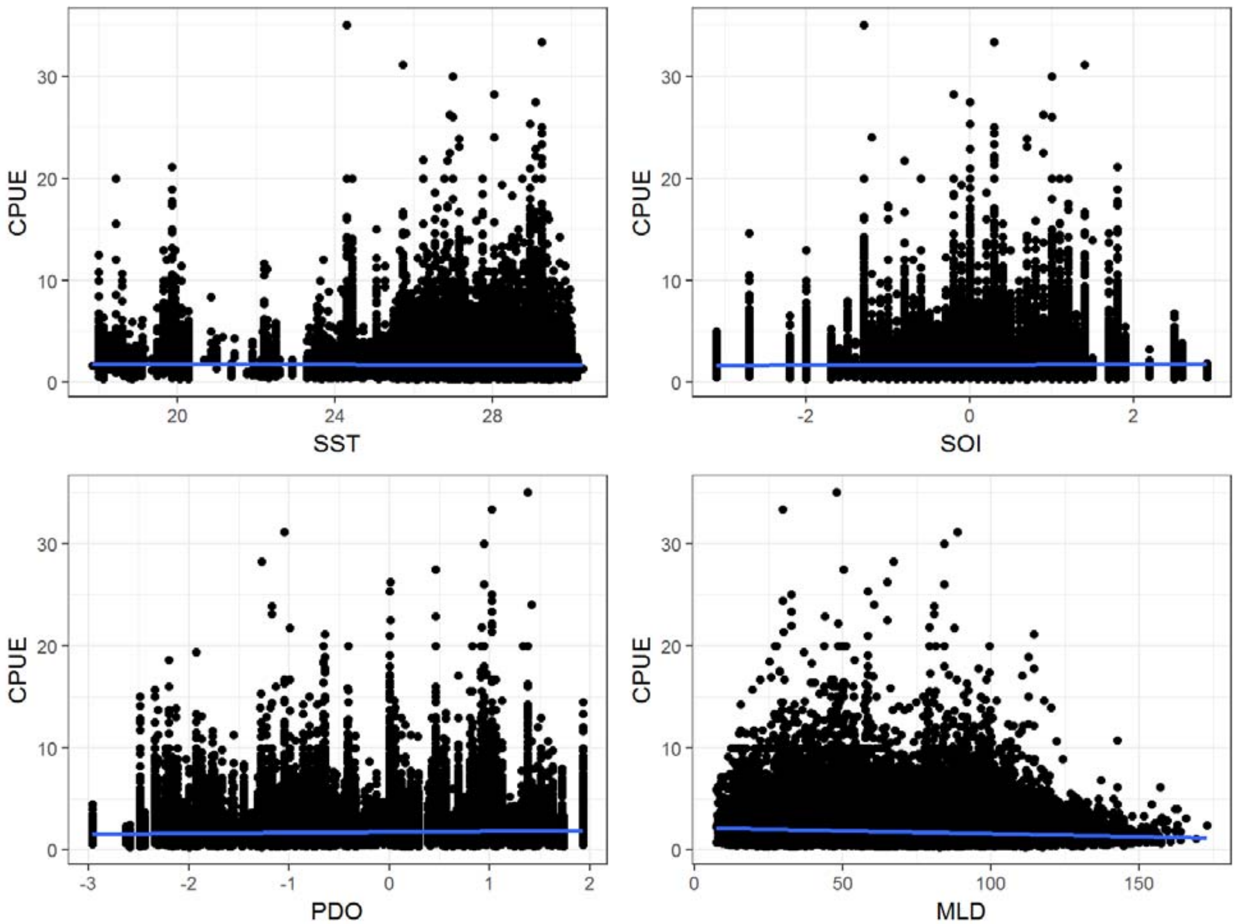

Figure 6. Positive combined CPUE versus the five environmental variables included in the analysis: Sea surface temperature (top left); Southern Oscillation Index (top right); Pacific Decadal Oscillation Index (bottom left); and mixed layer depth (bottom right). Blue line indicates a GAM smoother fit to the data.

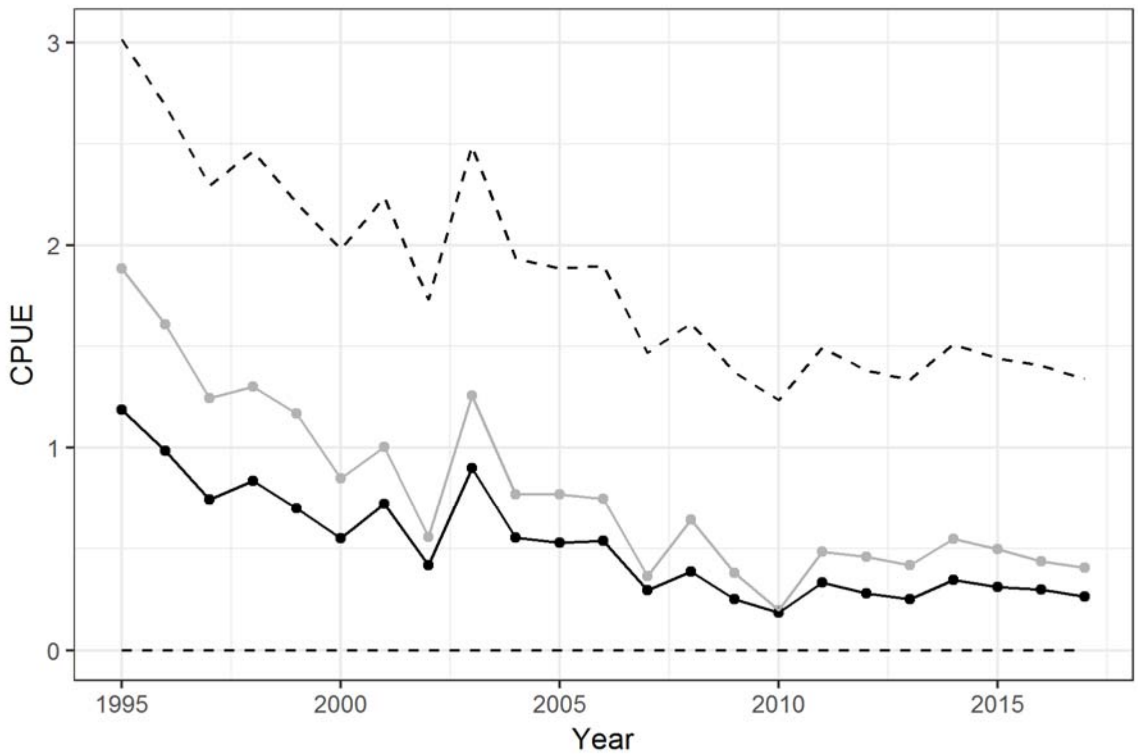

Figure 7. Nominal (grey) and standardized CPUE (black) with 95% confidence interval (black dashed line) around the standardized CPUE.

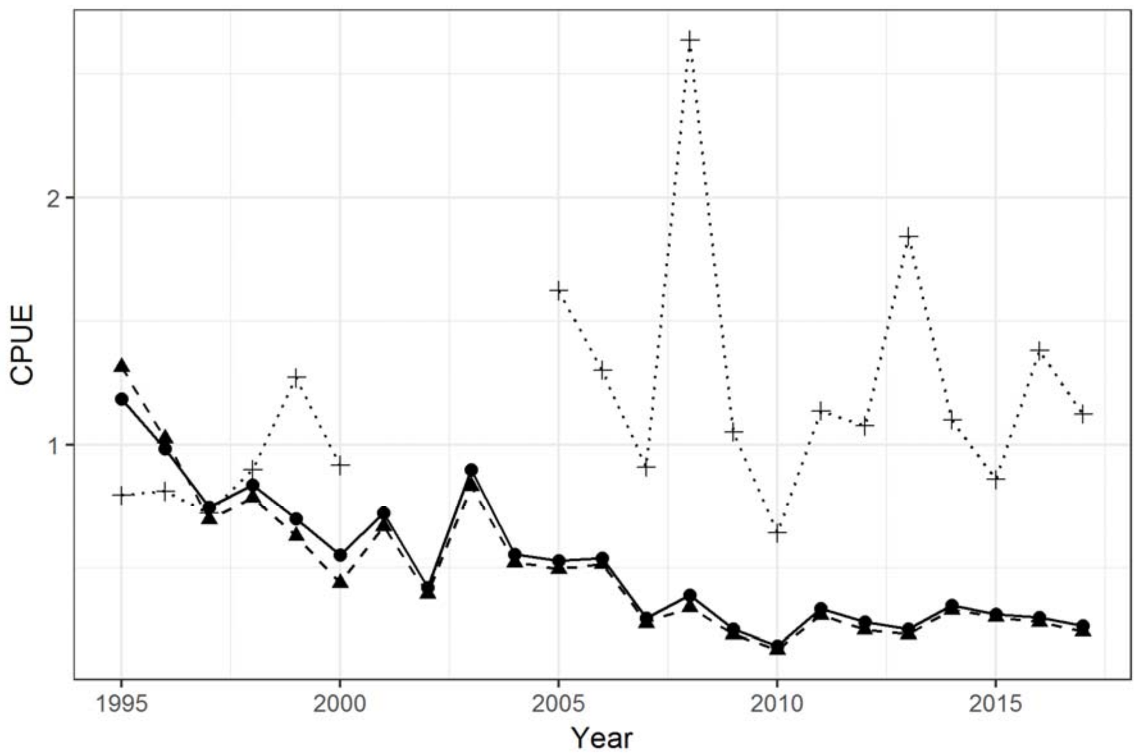

Figure 8. Standardized CPUE for combined data (solid line, circles); shallow-set only data (dotted line, crosses) and deep-set data (dashed line, triangles).

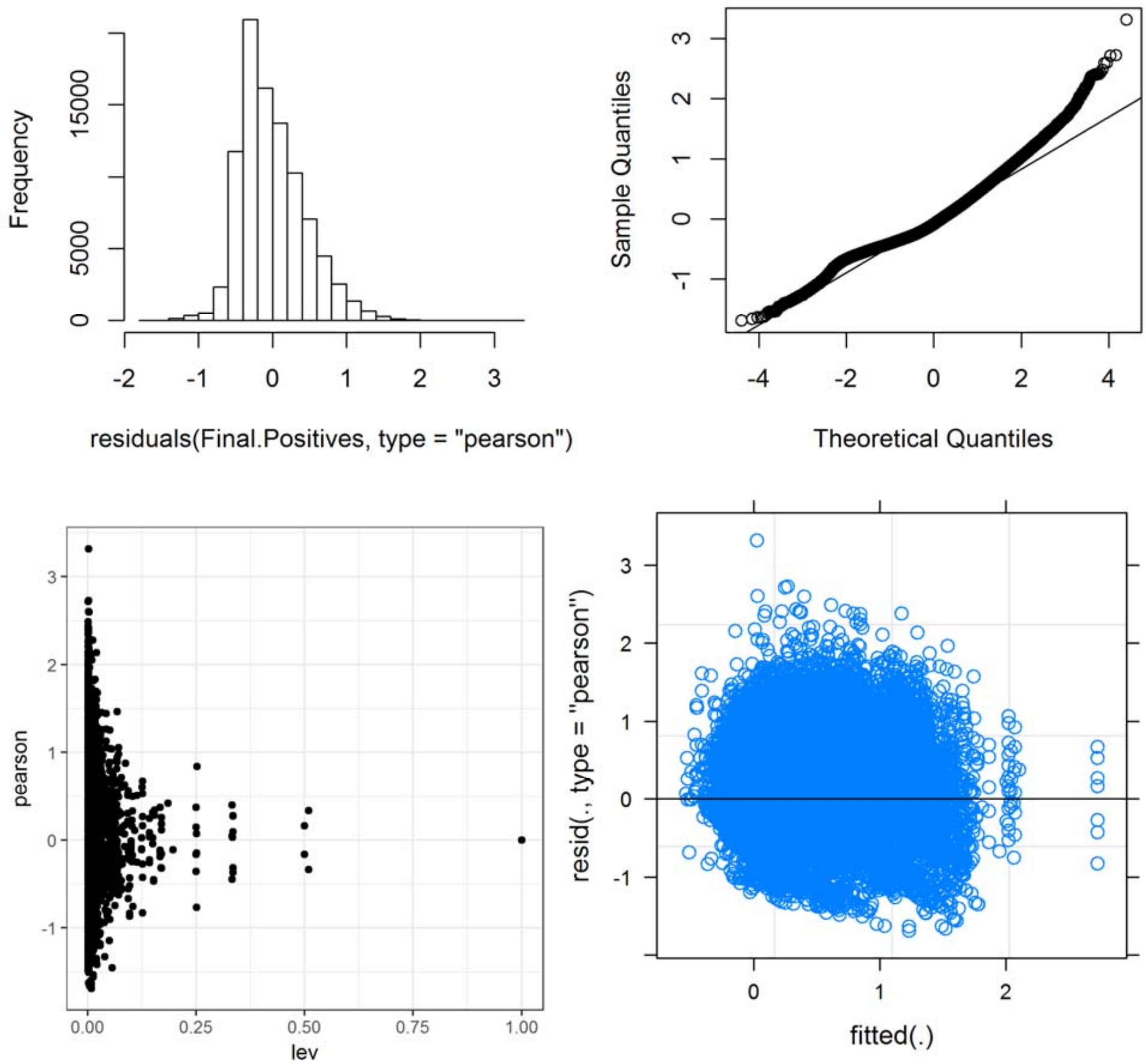

Figure 9. Diagnostic plots for positive catches: Histogram of standardized Pearson residuals (upper left) Normal Q-Q plot (upper right); Pearson residuals leverage plot (bottom left); Pearson residuals vs fitted values (bottom right).

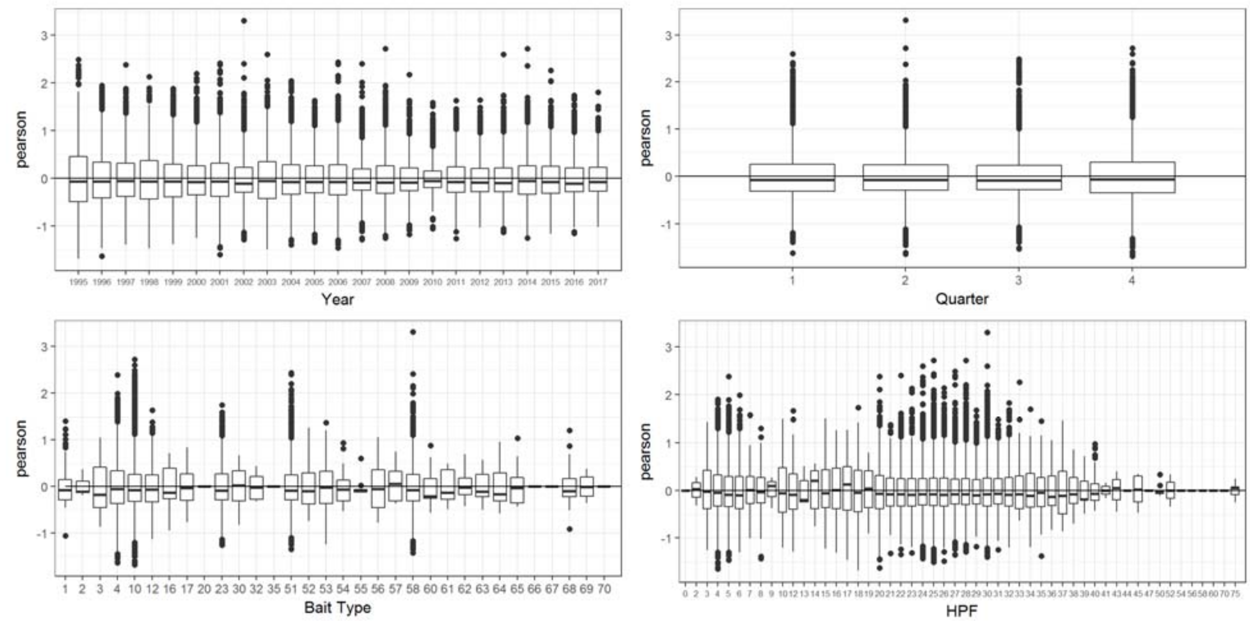

Figure 10. Residuals vs explanatory variables, positive catches: Year (upper left); Quarter (upper right); Bait type (lower left); and Hooks Per Float (lower right).

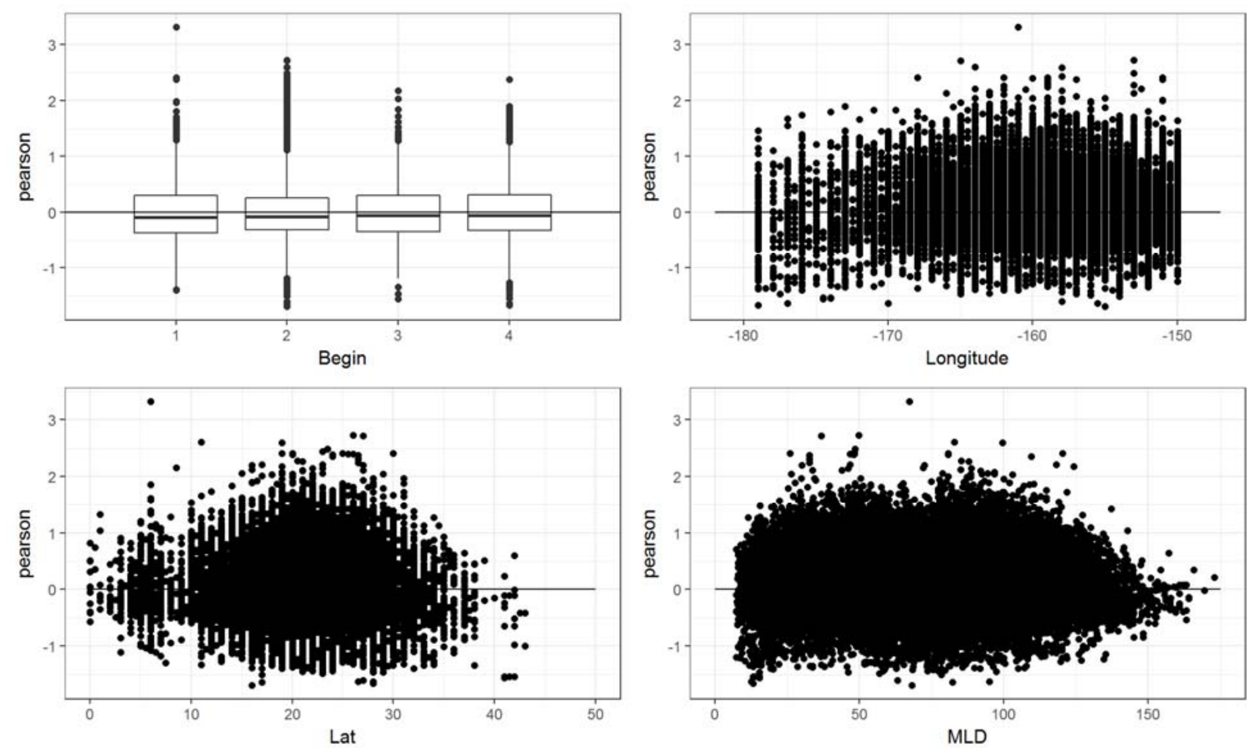

Figure 11. Residuals vs explanatory variables, positive catches: Begin (upper left); Longitude (upper right); latitude (lower left); and mixed layer depth (lower right).

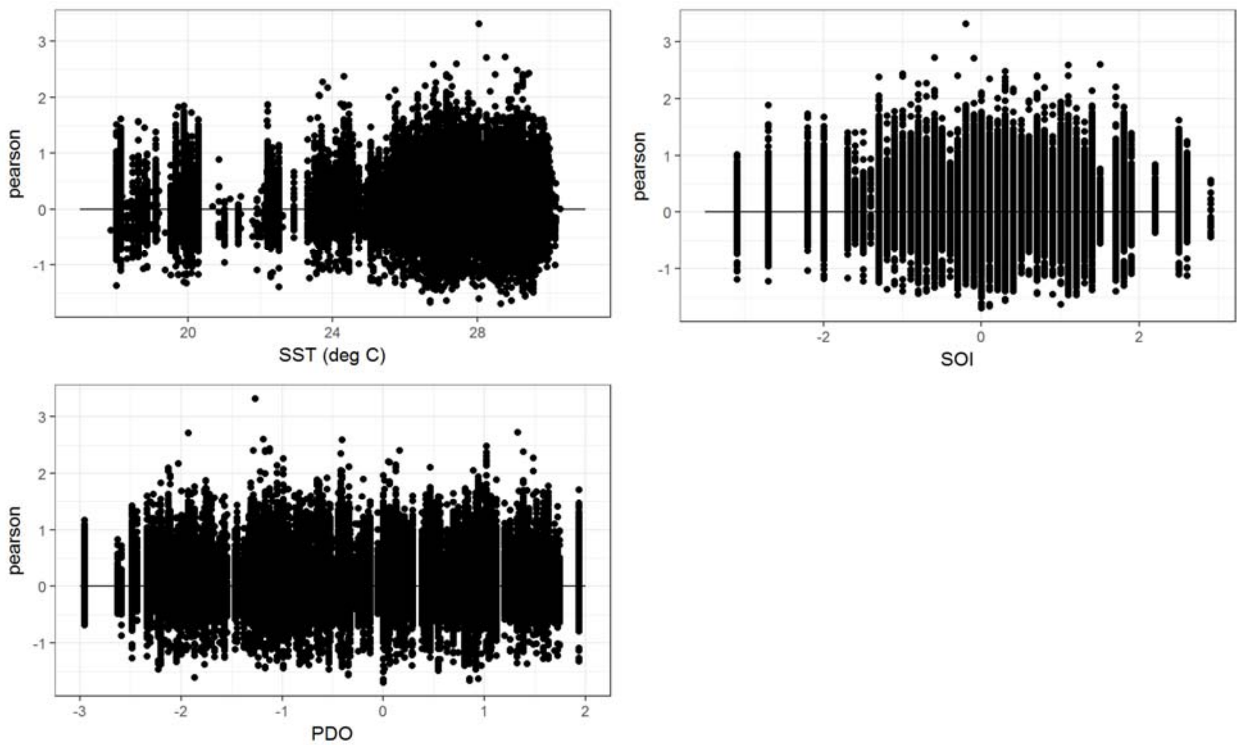

Figure 12. Residuals vs explanatory variables, positive catches: sea surface temperature (upper left); Southern Oscillation Index (upper right); and Pacific Decadal Oscillation Index (lower left).

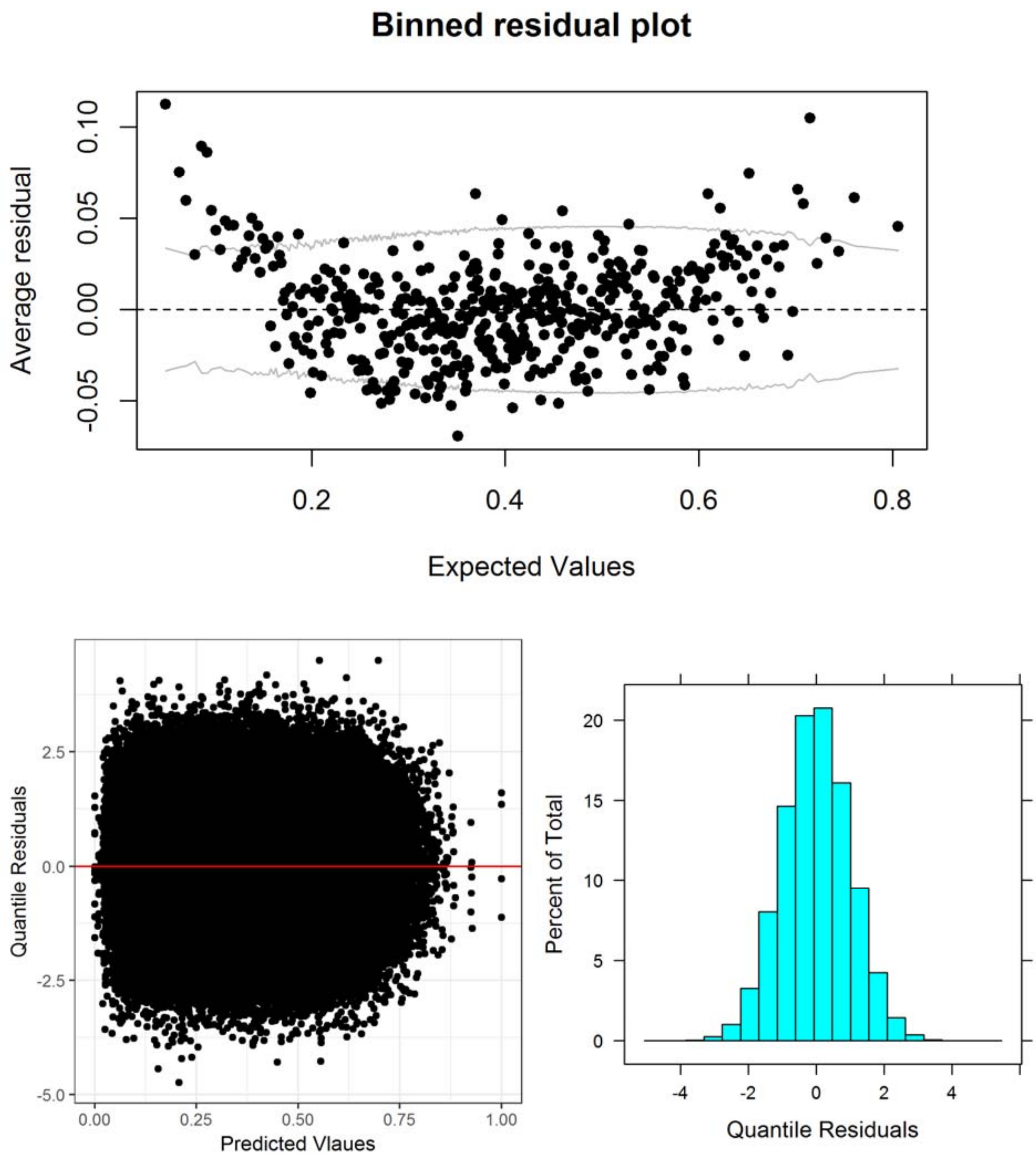

Figure 13. Diagnostics of the encounter probability binomial model: Binned residual plot (top); expected values vs quantile residuals (bottom left); and histogram of quantile residuals (bottom right).

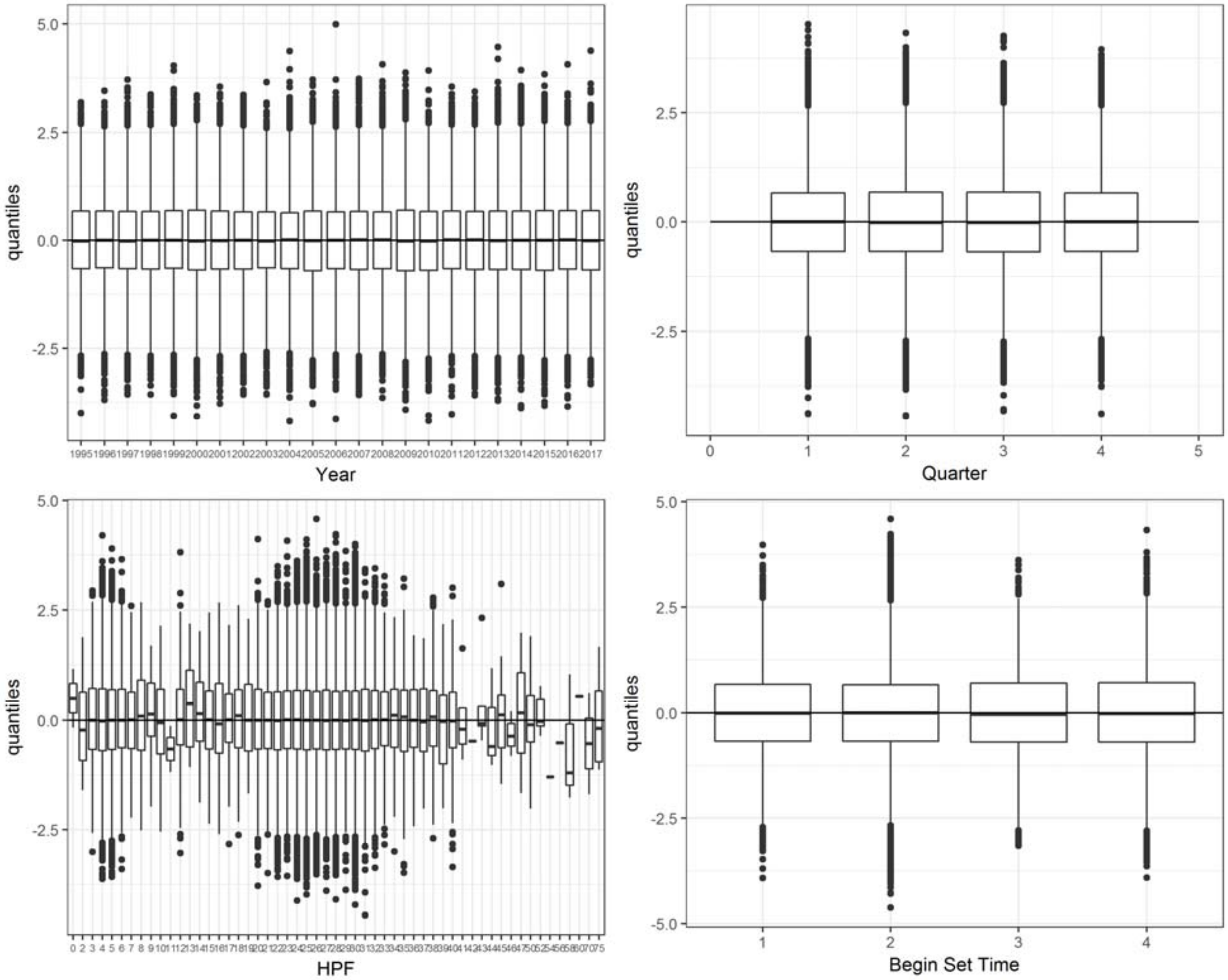

Figure 14. Quantile residuals vs explanatory variables for the encounter probability binomial model: Year (top left); quarter (top right); Hooks per float (bottom left); and Begin set time (bottom right).

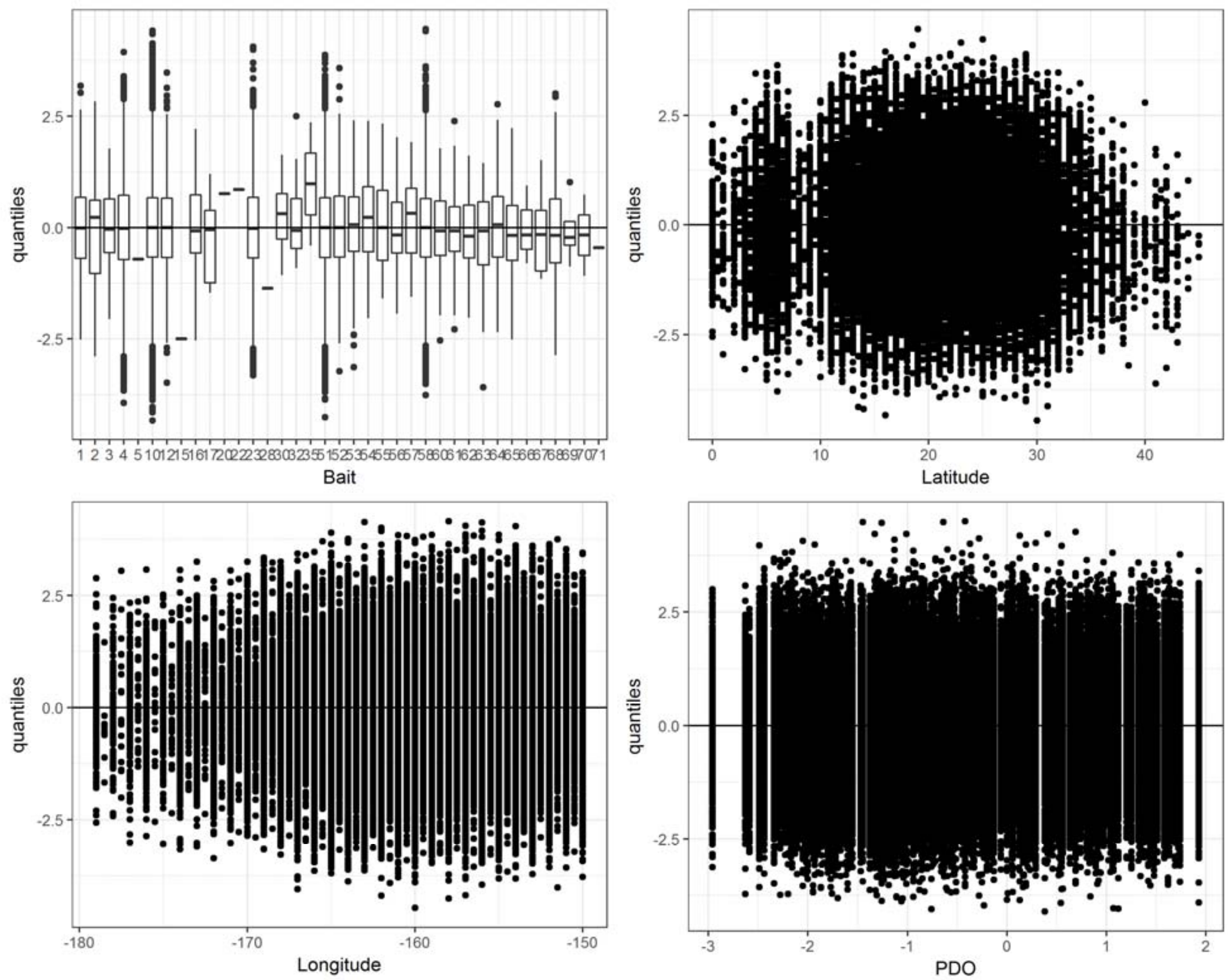

Figure 15. Quantile residuals vs explanatory variables for the encounter probability binomial model: Bait type (top left); Latitude (top right); Longitude (bottom left); and Pacific Decadal Oscillation Index (bottom right).

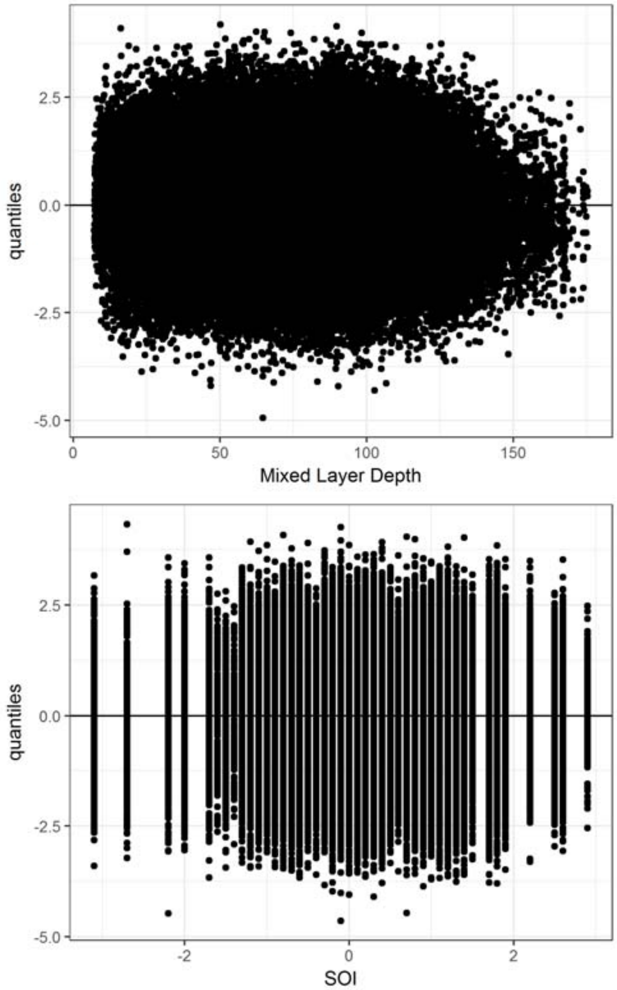

Figure 16. Quantile residuals versus explanatory variables for the encounter rates binomial model: Mixed layer depth (top) and Southern Oscillation Index (bottom).

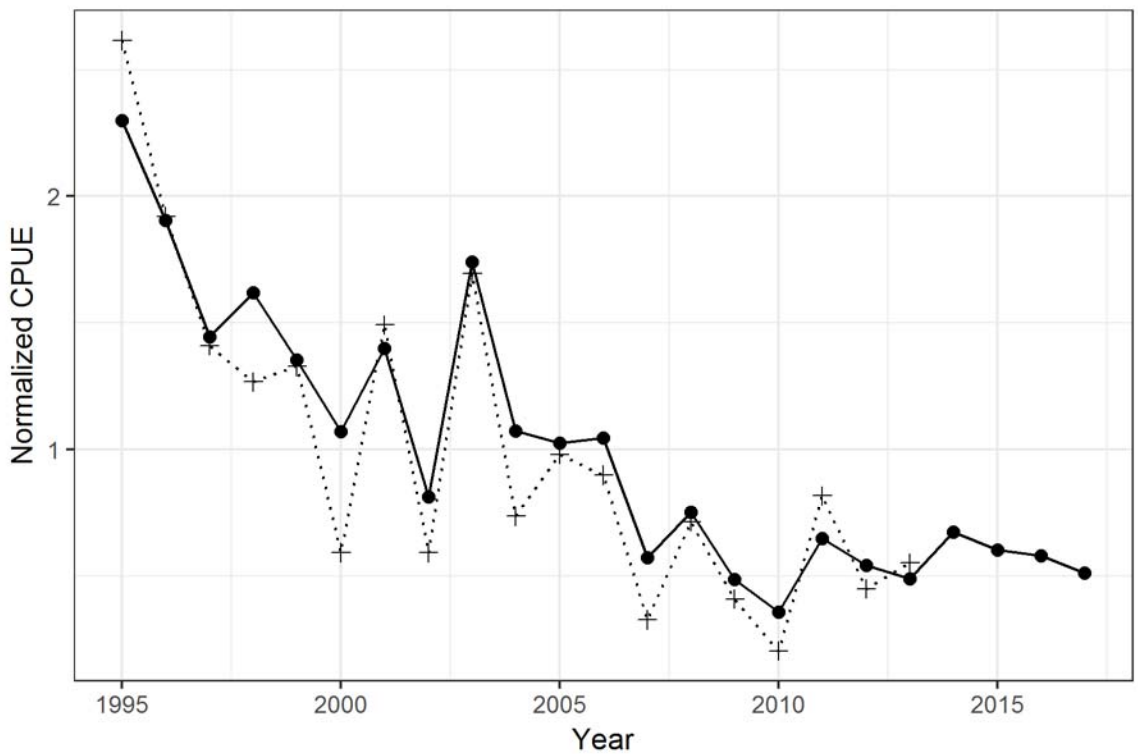

Figure 17. Normalized standardized CPUE from the 2015 assessment (dotted line, crosses) and the current 2019 assessment (solid line, circles).