A NEW APPROACH TO STANDARDIZE CATCH RATES FOR YELLOWFIN (*THUNNUS ALBACARES*) FOR THE SPANISH PURSE SEINE FLEET (1984-1995).

by

Soto M[1]., J. Morón[2] and P. Pallarés[1]

ABSTRACT

Standardized catch rates of yellowfin Indian stock using Delta-Lognormal method are presented. Models include information on factors affecting catchability (sonar, bird radar, net surface, speed and skipper experience) as well as abundance (year, area, season); floating school catch was also included as explanatory variable. Catch rate of free school yellowfin was used as response variable. Year and floating object catch were the most significant variables in all the approaches. GLM models seem not be appropriated to reflect the effect of the different equipment because the dynamic of the fishery in particular the lack of overlapping in the different characteristics range what makes the analysis unbalanced.

INTRODUCTION

This paper is a review of the document "Standardized catch rates for yellowfin (*Thunnus albacares*) for the Spanish purse seine fleet (1984-1995)" presented last year.

We use the same catch and effort data set to obtain an abundance index for yellowfin in the Indian Ocean for the period 1984-1995. We use a Delta-Lognormal method including information about boat's technical equipment and skipper during this period.

In this document we include some modifications in the variables used, in order to improve the previous analysis, trying to exploit the available information at maximum through the Delta-Lognormal approach.

MATERIAL Y METHODS

The data used in this new analysis are limited to the yellowfin catch in the areas where the school fishing is targeted, i.e. N-W Seychelles and S-E Seychelles. We do not consider the areas of Mozambique and Somali because the main target fishery is on log schools, neither Chagos because is an area for school fishing only developed during the last two years of the data series.

We use the data for the Spanish purse seine fleet operating in the Indian Ocean for the period 1984-1995. The information related to fishing power has been coded considering the date of introduction of the different equipment that increases detection and catching power of the boat.

The more relevant factors for the analysis of the fishing power increase are net , sonar, radar and boat speed. The following categories have been included:

Net was coded considering of the total size (T):

Net Category	Range
1	T< 0.3 km^2
2	0.3 km^2 =T < 0.4 km^2
3	0.4 km^2 =T < 0.5 km^2
4	T =0.5 km^2

For level 4 there were very few observations and was excluded from the analysis.

Radar was coded following the date of installation of every type of radar.

Radar Category	Radar Type
1	sin radar o radar 15
2	con radar 30
3	con radar 60.

The sonar classification was also coded by date of installation of the different sonar types.

Sonar Category	Sonar Type
1	without sonar
2	only one sonar, irrespective of the type
3	with two sonars (type 45 and 60)

The increase of boat speed was also taken in consideration as a factor for every boat individually. The average boat speed ranges from 10 to 15 knots.

A new variable to estimate skipper experience was created. Skippers were classified for the time they remain in the fishery since the beginning (1984) with the following scale:

Skipper Code	Time in the Fishery (months)
1	0-12
2	13-24
3	25-36
4	37-48
5	>48

[1] IEO. Corazón de María 8, 28002 Madrid Spain.
2OPAGAC. Ayala, 28006 Madrid

The zones selected for the analysis were coded as:

Area Code	Area
2	N-W Seychelles
4	E-S Seychelles

Every year was also divided in four fishing seasons with the following codes:

Season Code	Months
1	January-February
2	March-April-May-June
3	July-August-September-October
4	November-December

Catch information is available by set, position of the set and type of set (floating object or free school) by species: yellowfin, skipjack and bigeye. The effort information is available in number of sets (positive, nil and total), days at sea and fishing days but only by area and time strata, due to the characteristics of the fishey it is not possible split effort by type of set or species.

As response variable we considered the yellowfin catch rate on free school:

CPUE = YFT free school catch/total effort (fishing days)

The total floating objects (log) catch was introduced as explanatory variable defined as a proportion of the log catch over the total catch:

Log Catch= total floating object catch(yft+skj+bet+others) / total catch

In order to avoid the excess of zero catch, all the observations without information on some of the variables were eliminated.

Model 1

We would like to present a review of the Delta-Lognormal method used in the previous analysis to estimate relative abundance of yellowfin in the N-W Seychelles, Mozambique and E-S Seychelles. The method was applied to obtain three independent abundance indices for every zone. We will call this Model 1.

This model was divided in two components: the probability that cpue was bigger than zero, P{cpue>0}, and the distribution of the positive values of cpue. Both could be modelled independently to obtain, on one side, an adjustment of the positive cpue probability, and on the other hand, the expected cpue conditioned to obtain a cpue value bigger than zero. Then the Delta-Lognormal method comprises two lineal generalized models using the Bernoulli and Lognormal distributions respectively.

The relative cpue for every year in every of the three zones:

$CPUE_i = \mu_i p_i$, i = 1,2,3.

where m_i is a unit of standardized cpue for the positive catches in every zone, and p is a unit for the standardized proportion of positives.

To calculate a relative index for the whole zone, the three indices obtained were weighted by the number of observations in every zone, i.e.:

$CPUE_y = N_{y1}CPUE_{y1} + N_{y2}CPUE_{y2} + N_{y3}CPUE_{y3}$

where $CPUE_y$ is the yearly average cpue, and , N_{yi} is the number of observations in zone I for every year, and $CPUE_{yi}$ is the yearly average cpue in the year and zone i.

The following were considered the main factors: year, season, skipper, sonar, radar, net, speed and total log catch. The las factor is introduced in the analysis as

Log (totallog + K).

First order interactions between year, season and skipper with boat characteristics where also introduced.

The model to adjust cpue for positives is:

$Ln(CPUE) = X\alpha + Z\beta + \varepsilon$,

where cpue is the observations vector, X is the main factors matrix, a main factors parameter vector, Z is the interactions matrix, b is the interactions parameter vector and e is the independent error vector identically distributed that follow a $N(0, s^2)$.

On the other side, to estimate the proportion of positives, all the data set was used. A random variable Bernoulli type was created with value 0 or 1, depending if the cpue was nil or positive respectively. Then the average of this variable is calculated in every defined strata for every year, season, radar, sonar, net and speed combination, and the number of observations existing in every one is calculated.

The probability that the cpue could be positive, could be modelised through a Binomial GLM with the logit function as a connection between the explicative variables and the response variable, i.e., the appearanceof positive cpue is a Bernoulli random variable with a probability p given by:

$Log(p/(1-p)) = Z\alpha + Z\beta$, o bien,

$p = 1/ (1 + \exp\{ X\alpha + Z\beta \})$,

where X is the main factors matrix, a is the main factors parameters vector, Z is the interaction matrix, and b is the interactions parameters vector.

The following modifications were implemented in Model 1.

The analysis of the Indian Ocean data with the new variable “exskipper” instead of “skipper” is redone, that allow to include all the skippers in the analysis. In the previous case, it was necessary to select only those that remained in the fishery for more than six years, because otherwise there were levels with very few observations whose coefficients were difficult to estimate.

The variable “*log*” substitutes the one previously used “*log(totalobj +0.01)*”, that defined in this way is much more significant than before in all the other analysis.

The analytical approach of the variances of the Delta method estimated indices, that allows to include the upper and lower level in the cpue charts given by the standard annual error of the estimates.

In all the analysis only the interactions that have more sense were considered, i.e. those in which the skipper experience interact with the presence or absence of a particular equipment.

Starting from the initial model, a selection of the more significant variables is done, considering several criteria: the value of the statistics AIC, the proportion of the variability explained by the model, defined as *(Null Deviance – Residual Deviance)/ Null Deviance* and the F signification test and Chi-square.

Two kind of analysis were performed:

keeping the variable area and eliminating the variable season and calculating only an abundance index for zone 2 and 4 together,

- eliminating the variable area keeping the season, to eliminate the strong interaction existing among them, in a way that an index is estimated for every of the two zones and a global index is calculated for both areas is calculated, weighting every one by the square root of the observation numbers in every zone, to avoid weighting differently the area with more observations.

The Delta Lognormal method is applied in both analysis in the same way presented in Model 1.

Model 2

We will call Model 2 to that resulting of the application of the Delta-Lognormal method for every zone considered independently. The variable season is not considered in both zones, because by separating the analysis by areas, each of every one is associated to a season in which most of the catches are done. We take a look now to the analysis done for every zone.

AREA 2: N-W Seychelles

Area 2, N-W Seychelles is considered in this section. For the cpue>0 we started from the gaussian GLM.

log(cpuep) ~ year + expatron + net + speed + sonar + radar+ log + year:season + expatron:sonar + expatron:radar + expatron:net

the selection of variables was done based on the criteria previously pointed out, which results are shown in Table 1.

The last model was finally selected, which results of the ANOVA table for the error type III are shown in table 2.

The initial model for the proportion of positives was a binomial GLM with the logit function as link, and a selection of variables in the usual way was done, which results were shown in Table 3.

propor ~ year + expatron + net + speed+ sonar + radar + expatron:net+ expatron:sonar + expatron:radar

The last model was finally selected, which results of the ANOVA table for the error type III are shown in table 4.

Finally, to estimate the cpue of area N_W Seychelles by the Delta Lognormal Method, the adjusted annual average cpue in the gaussian GLM, previously reverting the logarithmical transformation, are multiplied by the annual average of P{CPUE>0} adjusted in the Binomial GLM. The Delta method is used to calculate the expected cpue variance, and it performs the Taylor development of the product to obtain an analytical expression of the variance.

Let us define x as positive cpue and y as proportion of positives. In principle we presume that both populations are independent and therefore the covariance is 0. For every year we define $m=E(x)$ and $p=E(y)$. The estimator of the expected cpue for every year is

$$g(x,y)= xy$$

and by the Taylor approximation, the variance estimator will be:

$$V[g(x,y)]= m^2*V(y)+ p^2*V(x)$$

We can see in figure 1 the expected cpue for area 2 with the upper and lower limits given by the annual standard error.

AREA 4: S-E SEYCHELLES.

Area 2, N-W Seychelles is considered in this section. For the cpue>0 we started from the gaussian GLM:

log(cpuep) ~ year + estacion + expatron + red + veloc + sonar + radar + objeto + year:estacion + expatron:sonar + expatron:radar + expatron:red

The model variables selection is done based on table 5.

The last model was finally selected, which results of the error type III ANOVA is shown in Table 6.

For the proportion of positives we start from the same model that in area 2. The selections of variables is done in the habitual way, which results are summarized in Table 7.

The last model was finally selected, which error type III ANOVA table results are given in Table 8.

We calculate the expected cpue in area 4 in the same way that in area 2, which graphical representation with the upper and lower levels given by the standard error every year is shown in Figure 2.

Once the index for every zone has been calculated, we weighted both indexes from zone 2 and 4 by the square root of the observations number in every zone and we calculate its variance by the Delta method.

Let us define x_1 as area 2 positive cpue, x_2 as area 4 positive cpue, y_1 area 2 proportion of positives, y_2 area 4 proportion of positives and $CPUE_2= E(x_1)$, $CPUE_4= E(x_2)$, $p_2= E(y_1)$, $p_4= E(y_2)$ for every year, and N_2 the number of observations in area 2 and N_4 the number of observations in area 4 for every year.

In principle we assume that all the four populations are independent and therefore its covariance is 0. The expected cpue estimator for every year is

$$g(x_1,y_1,x_2,y_2)= (N_2)^{1/2}*x_1y_1 + (N_4)^{1/2}*x_2y_2$$

and by the Taylor approximation, the variance of this estimator will be:

$$V[g(x_1,y_1,x_2,y_2)] = N_2*(V[x_1]*y_1^2 + V[y_1]*x_1^2) + N_4*(V[x_2]*y_2^2 + V[y_2]* x_2^2)$$

Model 3

We will call Model 3 to that resulting of applying the Delta-Lognormal Method to the data corresponding to the union of both zones, in a way that the area variable is eliminated from the analysis, but keeping the season variable.

As before, we introduce the variables Log and Expatron. In the same way, the Delta Lognormal method is applied, in which the cpue>0 is modelled through a gaussian model, while the proportion of observations with cpue>0 is modellized through the Binomial GLM.

For cpue>0 we start from the gaussian model.

log(cpuep) ~ year + season + expatron + net + speed + sonar + radar + log + year:season + expatron:sonar + expatron:radar + expatron:net

The selection of variables is done following the habitual criteria given in Table 9.

We finally select the last model, which variability proportion explained is *(1423.86-761.7049)/1423.86 = 0.4650423*, and the ANOVA table for the type III analysis is Table 10.

For the proportion of positives, the initial model was a Binomial GLM with the logit function as link, and a variable selection was performed on the habitual way, as we can see in Table 11:

propor ~ year + season + expatron + net + speed + sonar + radar + expatron:net + expatron:sonar + expatron:radar

The last model was selected, which explained variability proportion is *(1008.268-689.0685)/1008.268=0.316582*.

We calculate the expected cpue relative to 1986, and its variance by the Delta method.

RESULTS AND DISCUSSION

Model 2

AREA 2: N-W Seychelles.

Area 2, N-W Seychelles is considered in this section for the cpue>0 it started from the gaussian GLM.

Comparing with Model 1, in this Model, for the distribution of positives we can see that the F Statistic value corresponding to the log variable, is much bigger in this case *215.2428 >> 39.23202* . Therefore, although in both cases the log catch will be significant, defining the variable *log* as a proportion of log catch over the total catch, the sample shows a reject degree much bigger of the nil hypothesis that the variable will not be significant.

Besides, the variable year is also more significant that in Model 1, by eliminating from the analysis the variable *season*, the skipper experience (*Expatron*), is less significant than the variable *skipper*. Although, now the variable *net* is more significant replacing the variable *sonar*, as significant variable, as wel as the interaction among *Expatron* with *net*. Speed is slightly more significant in the last case.

The proportion of variability explained by the model *(840.2034-498.3111)/840.2034 = 0.4069161*, is bigger in this case that in Model 1, *(493.9918-310.121)/493.9918=0.3722143*.

In respect with the proportion of positives, removing *season* from the analysis, it gets worst because the only variable appearing as significant is *year*. Variable *log* is not included because is continuous and we will have to categorize it, to be able to calculate the proportion of observations of cpue>0 that will exist for every category. Maybe in a following analysis, we could test if the model improves with that option, but for the time being, this variable was not included.

The proportion of variability explained by the model for the proportion of positives is *(431.7917-88.1492)/431.7917=0.101073*, while in Model 1 we were getting *(606.1353-515.1864)/606.1353=0.1500472*.

AREA 4: S-E SEYCHELLES.

In the same way that in Area 2, the variable *log* is much more significant defined in this way as the log catch proportion, because the F statistic value is bigger that in the original document model *108.52 > 2.02027*. Equally, *year* is slightly more significant than before, but *sonar* and *radar* disappear as important variables, in exchange the *net* and *speed* have a bigger signification. The variability proportion explained by the model in this case is *(382-365)/382 = 0.04450262*, much lower that in the previous case mucho *(370.6122-168.2984)/370.6122 = 0.5458908*, because by removing the variable *season*, a lot of information is lost.

For the proportion of positives, the analysis does not improve by eliminating *season*. Although the variable *speed* is not very significant, it was finally included in the model to isolate, in a way, the *year* effect. The proportion of variability explained by the model *(374.3623-289.0467)/374.3623 = 0.2278958*, is less than in model 1 *(482.8845-219.5806)/482.8845 = 0.545273*, because the variable *season* was very significant.

Model 3

We can see in Table 10, that when the number of observations is bigger, the model works better, the variable *log* continue to be most significant by far, while *year* and *season*, and the interaction among them are also significant. The skipper experience does not appear as significant, neither the *radar*, while *net* and *speed* are significant and to a lesser extent *sonar*.

For the proportion of positives, the analysis can not be improved much either, because the only variables appearing as significant are *year* and *season* and its interaction.

In Table 12, expected cpue results related to 1986 are shown form the three models.

In figures 3.a, 3.b and 3.c, we can see the expected cpue results related to 1986 in model 1,2 and 3, respectively. The upper and lower limits given by the standard error from the models 2 and 3, are shown in figures 3.b and 3.c.

Comparing the expected cpue in the three analysis, we can graphically see that there is a very similar trend in the three cases. However, the relative values, and the confidence intervals vary among models 2 and 3.

In Model 2, when variable *season* is not considered and the index for area 2 and 4 is weighted by the square root of the number of observations in every area, a relative expected cpue bigger than in Model 3 is obtained for all the years. However, the standard errors of the expected cpue for every year are much lower in Model 3, specially for the years where cpue is maximum.

The obtained results in Model 2, are very similar to the initial Model 1, although for almost all the years, the expected cpue estimated by Model 1 is bigger.

Therefore, there are differences between the two models, while there is not a clear criteria to determine which is more adequate. On one side, the two zones have been separated disregarding the variable *season*, and the results are very similar to the initial model in which three areas where considered and the variable *season* was included. Therefore the variable *season* in this context is not important to determine the index.

Eliminating area3 (Mozambique), the estimated cpue trend, was not affected because it is correlated to a very small data group And the general trend is dominated by the effects of areas 2 and 4 (NW and SE Seychelles).

In model 2 for areas 2 and 4, and Model 3, the variables finally selected in the gaussian GLM applied to the distribution of observations with positive cpue, are not the same. The significant variables, common to the three models are *year*, *net*, *speed* and *log*. Let observe how the obtained coefficients in every of the three GLM for each of the variables are.

In Table 13, the coefficients corresponding to the variable *year* were obtained in the gaussian GLM for cpue>0 in model 3, *(posit2year)*, and in model 2 for zone 2 *(a2posit2year)* and zone 4, *(a4posit2year)*.

Its graphic representation is given in Figure 4.

Table 14 shows the corresponding coefficients to the variable *net* obtained in the gaussian GLM models for cpue>0 in model 3 *(posit2net)*, and in the model 2 for zone 2 *(a2posit2red)*, and zone 4 *(a4posit2red)*.

Its graphic representation is given in Figure 5.

In Table 15 coefficients corresponding to the variable speed obtained in the gaussian GLM for cpue>0 in model 3 *(posit2speed)*, and in model 2 for zone 2 *(a2posit2speed)*, and the zone 4 *(a4posit2speed)*.

Its graphic representation is given in Figure6.

All the coefficients have a first level value 0, because by applying the gaussian GLM model, the *contrast treatment* option of the S-Plus package was used, that corresponds to assign value zero to the first coefficient, in a way that the others appear related to this level.

The continuous variable *log* has the following coefficients:

Model 3: *-1.748431*

Model 2, area 2: *-1.854503734*

Model 2, area 4: *-1.78194746*

In all the cases it has negative values, as could be expected, because the dependent variable is related to the free school catch, that it is complete with the log catch. The coefficient values are very similar in all the three cases, and in all the models, this is the more significant variable.

The coefficients of variable *year* have a similar trend for Model 2 in area 4 and Model 3. However, Model 2 for area 4 has all coefficients bigger that in Model 2. Results from this two models do not vary much from those of Model 1.

Model 2 for area 2 has different trend from the others, and variable *year* presents coefficients with less variability.

Intuitively, so much for the *net* as for the *speed*, the expected coefficients should increase according to every factor level, since expected yields should be bigger when the boat has a bigger net and her speed increases. However the coefficients do not show that clearly.

In Model 2 for area 4, if increasing coefficients for *net* are obtained, for the other two models the results are not intuitive.

In Model 2 for area 2 s*peed* has increasing coefficients from 12 knots and for Model 3 from 11 knots. In Model 2 for area 4, coefficients vary with a clear trend. The value for *speed* coefficients in the three models for levels above 15 knots is very similar.

However, at the time of estimating the different abundance index the importance of the different equipment, is not clearly reflected in GLM models, since the significance of a certain boat characteristic, vary much from one model to the other, depending on the number of observations and the zone considered, being in certain occasions the results contrary to logical intuition.

The lack of overlapping in the different characteristics range makes the analysis difficult, since most of the boats introduce the new equipment at the same time, it is difficult to compare along the period, the yields when the equipment has been improved.

On the other side, the considered period is very limited to be able to perform a good analysis, due to the difficulty in compiling this kind of information.

REFERENCES

AITCHISON, J., 1955: On the distribution of a positive random variable having a discrete probability mass at the origin. Journal of the American statistical association, 50: 901-908.

BISHOP, J., WANG, Y. G. Y DIE. D., 2000: A generalised estimating equations approach for analysis of the impact of the new technology on a trawl fishery. Australia New Zealand Journal of statistics.

COOKE, J. G, K. LANKESTER, 1977: Consideration of statistical models for catch-effort indices for use in tuning VPA's. SCRS/95/77.

COOKE, J. G., 1996: A procedure for using catch-effort indices in bluefin tuna assessments. SCRS/96/63.

DRAPER, N. R., Y SMITH, H., 1981: Applied reggression analysis. Ed. Wiley series in probability and mathematical statistics, second ed.

GAVARIS, S., 1980: Use a multiplicative model to estimate catch rate and effort from commercial data. Can. J. Fish. Aquat. Sci. 37: 2272-2275.

HOEY, J., CONSER, R Y DUFFIE, E., 1988: Catch per unit of effort information from the U.S. swordfish fishery. ICCAT working document. SCRS/88/21.

HSU, C. C. Y LIU, H. C., 1999: The updated catch per unit of effort of bigeye tuna for Taiwanese longline fishery in the Atlantic. SCRS/99/130.

LALOE, F., 1998: Model identification for flexible multifllet-multispecies fisheries. A simulation study. Fish. Res., 37: 193-202.

LARGE, P. A., 1992: Use a multiplicative model to estimate relative abundance from commercial CPU data. ICES J. Mar. Sci., 49: 253-261.

LO, N. C. H., N., L. D. JACOBSON Y L. J. SQUIRE, 1992: Indices of relative abundance from fish spotter data based on Delta-Lognormal models. Can. J. Fish. Aquat. Sci. 49: 2515-2526.

MCCULLAGH, P. AND NELDER, J. A., 1989: Generalized linear models, second ed. Ed. Chapman & Hall.

MEYERS, R. A. Y PEPIN, P., 1990: Biometrics, 46: 1185-1192.

MILLISCHER, L. Y GASCUEL, D., 1998: Individual based modeling of fishing tactics. ICES anual conference.

MIYABE, N., 1988: Estimation of standarized cpue for the Atlantic swordfish using the data from the Japanese logline fishery. SCRS/88/20.

MIYABE, N., 1991: Trend of cpue for Atlantic swordfish caught by the japanese logline fishery in the Atlantic ocean. SCRS/91/34.

NEYMAN, K., 1997: Bayesian averaging of generalized linear models for passive integrated trnsponder tag recoveries from salmonids in the Snake river. N. Am. J. Fish. Manage. 17: 362-377.

O'BRIEN, C.M., KELL, L. T., SANTIAGO, J. Y ORTIZ DE ZÁRATE, V., 1997: The use of generalized linear models for the modelling of catch-effort series. II. Aplicaton to north Atlantic albacore surface fishery. ICCAT working document. SCRS/97/49.

O'BRIEN, L. Y RAGO, P., 1996: An application of the generalized additive model to grounfish survey data with atlantic cod off the northeast coast of the United States as an example. NAFO Sci. Coun. Studies, 28: 79-95.

ORTEGA-GARCÍA, S. Y GÓMEZ-MUÑOZ, V., 1992: Standarization of fishing effort using principal component analysis of vessel characteristics: the Mexican purse-seiners. Sci. Mar., 56: 17-20.

ORTIZ, M., TURNER, S. C. Y BROWN, C. A., 1998: Standarized catch rates for small bluefin tuna, *Tunnus thynnus*, from the reel fishery off the northeast United States from 1980-1997. ICCAT working document. SCRS/98/59.

PENNINGTON, M., ON TESTING OF LOGNORMAL-BASED MODELS, 1971: Biometrics, 47: 1623-1624.

PENNINGTON, M., 1996: Estimating the mean and variance from highly skewed marine data. Fishery Bulletin 94:498-505.

PEÑA, D., 1998: Etadística, modelos y métodos. Vol. 1 y 2. Ed. Alianza Editorial.

PUNSLY, R.: Estimation of the relative annual abundance of yellowfin tuna, *Thunnus Albacares*n in the easternPacific ocean durin 1970-1985.

PUNT, A. E., WALKER, T. I., TAYLOR, B. L. Y PRIBAC, F., 1999: Standarization of catch and effort data in a spatially-structured shark fishery. Fish. Res. 45: 129-145.

ROBINS, C. M., WANG, Y. G. Y DIE. D., 1999: The impact of global psitioning systems and plotters on fishing power in the northern prawn fishery, Australia. Canad. J. Fish. Aquat. Sci., 55: 1645-1651.

ROBSON, D. S., 1966: Estimation of the relative abbundance fishing power of individual ships. ICNAF Res. Bul., 3: 5-14.

SMITH, S. J., 1988: Evaluating the efficiency of the delta-distribution mean estimator. Biometrics, 44: 485-493.

SWARTZMAN, 1992: Spatial analysis of bearing sea groundfish survey data usin generalised additive models. Can. J. Fish. Aquat. Sci, 49: 1366-1378.

SWARTZMAN, G., SILVERMAN, E., WILLIAMSON, N., 1995: Relating trends in walleye pollock (Theragra chalcogramma) abundance in the Bearing sea to environmental factors. Can. J. Fish. Aquat. Sci. J. Can. Sci. Halieut. Aquat., 52: 369-380.

STEFÁNSSON, G., 1996: Analysis of groundfish survey abundance data: combining the GLM and delta approaches. ICES Journal of Marine Science, 53: 577-588.

WADE, P. R., 1999: A comparison of statistical methods for fitting poplation models to data. Marine mammal survey and assessment methods, Garner et al. (eds), 249-270.

Table 1.- Results from the ANOVA and Akaike statistic value from the nested gaussian models for area 2. Y=Year, Ex=Expatron, Re=Net, S=Sonar, Ra=radar, V=Speed, O=Log.

Model	Variation source	d.f.	Null Dev.	Res. Dev.	AIC
Y+Ex+Re+S+Ra+V+O	Model	25	840.2034	510.3538	559.7735
	Error	537			
	Total	562			
Y+Ex+Re+S+Ra+V+O+Ex:Re+Ex:So +Ex:Ra	Model	47	840.2034	476.5370	565.3673
	Error	515			
	Total	562			
Y+Ex+Re+V+O+Ex:Re	Model	30	840.2034	498.3111	556.3849
	Error	532			
	Total	562			

Table 2: ANOVA type III results from the gaussian model selected for area 2.

	Df	**Sum of Sq**	**Mean Sq**	**F Value**	**Pr(F)**
year	9	54.5770	6.0641	7.6277	0.0000000
expatron	4	5.9738	1.4934	1.8785	0.1127946
red	2	6.9041	3.4520	4.3421	0.0134732
veloc	5	8.8744	1.7749	2.2325	0.0498536
radar	2	5.7325	2.8662	3.6053	0.0278468
objeto	1	171.1210	171.1210	215.2428	0.0000000
expatron:red	7	12.6280	1.8040	2.2691	0.0277531
Residuals	532	422.9474	0.7950		

Table 3: ANOVA results and Akaike statistic value of the binomial model nested in area 2.

Model	Variation source	d.f.	Null Dev.	Res. Dev.	AIC
Y+Ex+Re+S+Ra+V	Model	24	431.7917	375.3453	461.0406
	Error	219			
	Total	243			
Y+Ex+Re+S+Ra+V+Ex:Re+Ex:S+E x: Ra	Model	46	431.7917	337.4472	498.4626
	Error	197			
	Total	243			
Y	Model	9	431.7917	388.1492	421.3243
	Error	234			
	Total	243			

Table 4: ANOVA type III results from the gaussian model selected for area 2.

	Df	**Sum of Sq**	**Mean Sq**	**F Value**	**Pr(F)**
year	9	34.6142	3.846026	2.898477	0.002865467
Residuals	234	310.4976	1.326913		

Table 5: ANOVA results and Akaike statistic values from the gaussian models nested for area 4. Y=Year, Ex=Expatron, Re=net, S=Sonar, Ra=radar, V=speed, O=log.

Model	Source of variation	d.f.	Null Dev.	Res. Dev.	AIC
Y+Ex+Re+S+Ra+V+O	Model	25	576.5284	323.5141	370.6366
	Error	357			
	Total	382			
Y+Ex+Re+S+Ra+V+O+Ex:Re+ Ex:S+Ex:Ra	Model	47	576.5284	309.4133370	398.0810
	Error	335			
	Total	382			
Y+Re+V+O	Model	17	576.528434	327.3611	359.6488
	Error	365			
	Total	382			

Table 6: Error type III ANOVA results from the gaussian model selected for area 4.

	Df	Sum of Sq	Mean Sq	F Value	Pr(F)
year	9	87.9634	9.77371	12.8214	.00000000
net	2	3.9947	1.99737	2.6202	0.07415734
speed	5	8.9139	1.78277	2.3387	0.04138702
log	1	82.7293	82.72929	108.5261	0.00000000
Residuals	365	278.2388	0.76230		

Table 7: ANOVA results and Akiake statistic results from the binomial model nested in area 4.

Model	Source of variation	d.f.	Null Dev.	Res. Dev.	AIC
Y+Ex+Re+S+Ra+V	Model	24	374.3623	276.5183	349.2863
	Error	190			
	Total	214			
Y+Ex+Re+S+Ra+V+Ex:Re+Ex:S+Ex:Ra	Model	46	374.3623	248.4579	387.4760
	Error	168			
	Total	214			
Y+V	Model	14	374.3623	289.0467	332.4037
	Error	200			
	Total	214			

Table 8: Error type III ANOVA results from the binomial model selected for area 4.

1	Df	Sum of Sq	Mean Sq	F Value	Pr(F)
year	9	58.6864	6.520708	5.802929	0.00000036
speed	5	10.9796	2.195930	.954209	0.08704089
Residuals	200	224.7385	1.123692		

Table 9: ANOVA results and Akaike statistic value from the gaussian models nested by the union of area 2 and 4. Y=Year, E=Season, Ex=Expatron, Re=Net, S=Sonar, Ra=radar, V=Speed, O=Log.

Model	Variation Source	d.f.	Null Dev.	Res. Dev.	*AIC*
Y+E+Ex+Re+S+Ra+V+O	Model	28	1423.860	867.6645	922.5440
	Error	917			
	Total	945			
Y+E+Ex+Re+S+Ra+V+O+Y:E+Ex:Re+Ex:S+Ex:Ra	Model	77	1423.860	732.1296	863.7105
	Error	868			
	Total	945			
Y+E+Ex+Re+S+V+O+	Model	49	1423.860	761.7049	846.7166
Y:E	Error	896			
	Total	945			

Table 10: Type III ANOVA results from the gaussian model selected for the union of area 2 and 4.

	Df	**Sum of Sq**	**Mean Sq**	**F Value**	**Pr(F)**
year	9	106.3560	11.8173	16.3334	0.00000000
estacion	3	38.6935	12.8978	17.8268	.00000000
red	2	7.9720	3.9860	5.5093	.00418749
veloc	5	13.0201	2.6040	3.5991	.00314417
sonar	2	3.8035	1.9018	2.6285	.07274086
objeto	1	253.3119	253.3119	350.1160	0.00000000
year:estacion	27	107.8265	3.9936	5.5197	0.00000000
Residuals	896	648.2635	0.7235		

Table 11: ANOVA results and Akaike statistic value from the binomial models nested for the union of area 2 and 4.

Model		d.f.	Null Dev.	Res. Dev.	AIC
Y+E+Ex+Re+S+Ra+V	Model	27	1008.268	781.3035	854.9618
	Error	594			
	Total	621			
Y+E+Ex+Re+S+Ra+V+	Model	76	1008.268	635.1413	814.6124
Y:E+Ex:Re+Ex:S+Ex:Ra	Error	545			
	Total	621			
Y+E+S+Y:E	Model	41	1008.268	684.8161	783.9964
	Error	580			
	Total	621			
Y+E+Y:E	Model	39	1008.268	689.0685	783.7858
	Error	582			
	Total	621			

Table 12: Expected cpue related to 1986 for the model 1,2 and 3.

Year	Model 1 cpue	Model 2 cpue	Model 3 cpue
1986	1.00	1.00	1.00
1987	1.36	1.07	0.83
1988	4.20	4.01	2.39
1989	1.77	1.19	0.55
1990	3.09	3.03	1.73
1991	5.80	4.87	2.18
1992	3.62	3.12	1.66
1993	1.48	1.86	1.25
1994	2.47	2.82	1.95
1995	1.56	1.29	0.68

Table 13: Year coefficients form the different models.

Year	Model 3	Model 2 zone 2	Model 2 zone 4
86	0.00	0.00	0.00
87	-0.38	0.00	-0.43
88	-0.32	-0.18	0.23
89	-0.64	-1.03	-0.33
90	-0.62	-0.39	0.34
91	0.42	-0.29	0.84
92	-0.39	-0.29	-0.09
93	-0.81	-0.39	0.06
94	0.22	-0.58	0.43
95	-1.15	-0.45	-0.30

Table 14: Net coefficients form the different models.

net	Model 3	Model 2 zone 2	Model 2 zone 4
red1	0.00	0.00	0.00
red2	-1.54	0.04	-0.12
red3	-0.78	0.22	0.11

Table 15: Speed coefficients form the different models.

speed	Model 3	Model 2 zone 2	Model 2 zone 4
10	0.00	0.00	0.00
11	0.43	0.56	-0.12
12	-0.17	0.22	0.11
13	0.01	0.31	0.19
14	0.40	0.13	0.15
15	0.45	0.50	0.55

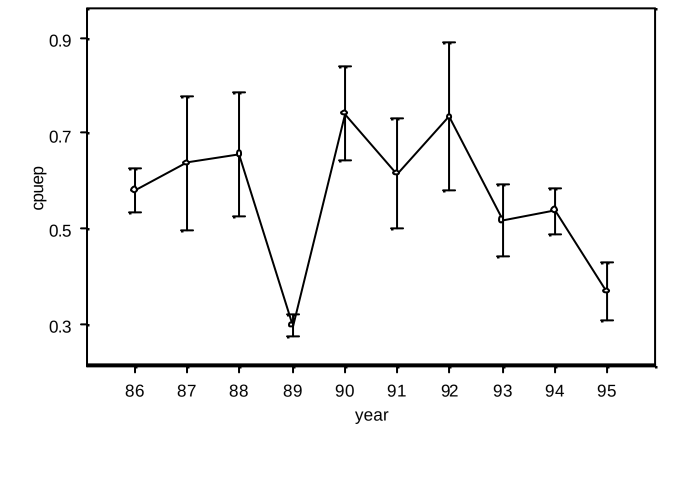

Figure 1: Expected cpue for area 2 with the upper and lower limits given by the annual standard error.

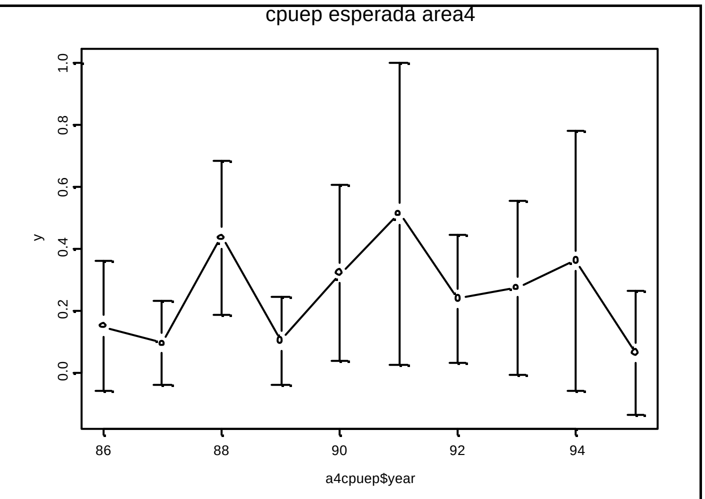

Figure 2: Expected cpue for area 4 with upper and lower limits given by the annual standard error.

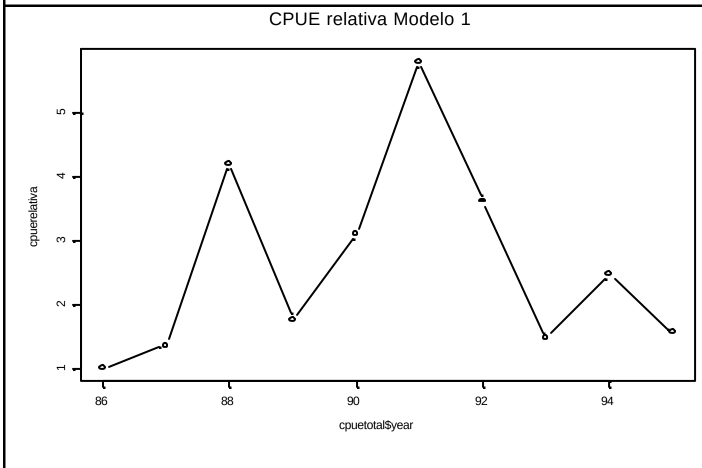

Figure 3.a: Expected cpue related to 1986 from model 1

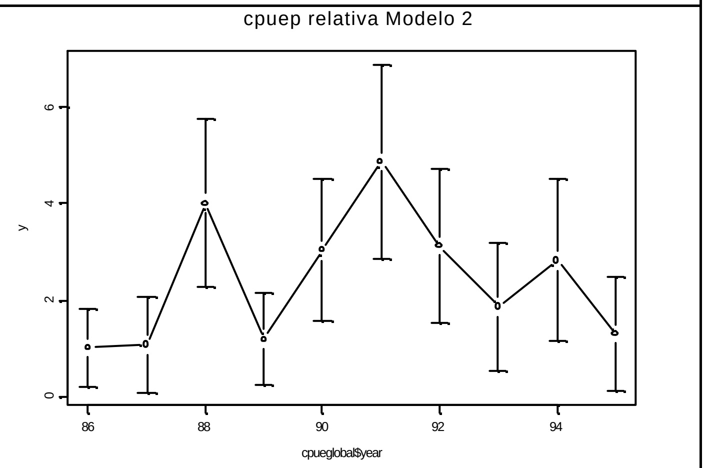

Figure 3.b: Expected cpue related to 1986 from model 2.

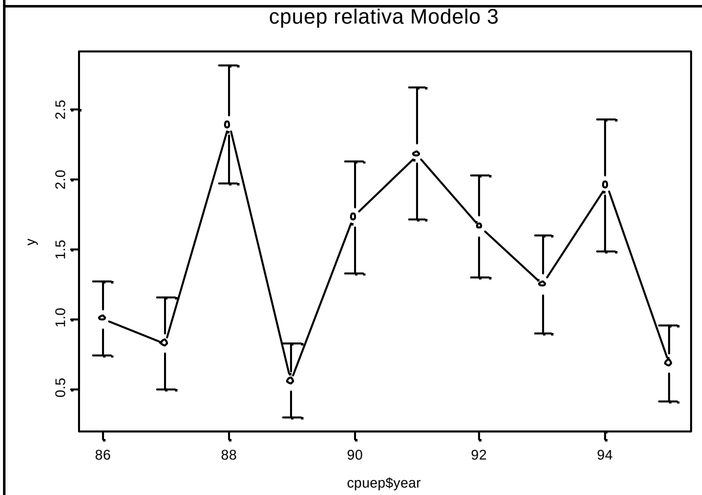

Figure 3.c: Expected cpue related to 1986 from model 3.

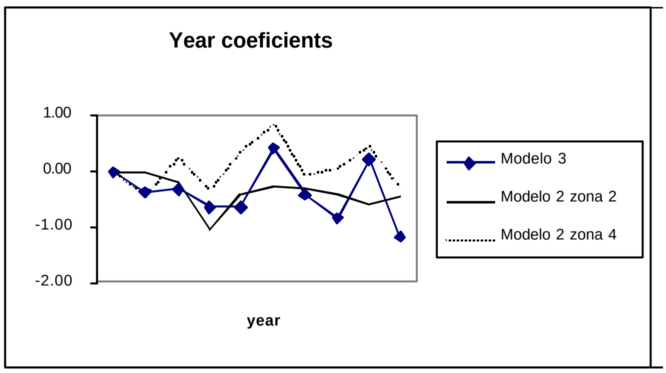

Figure 4: Year coefficients

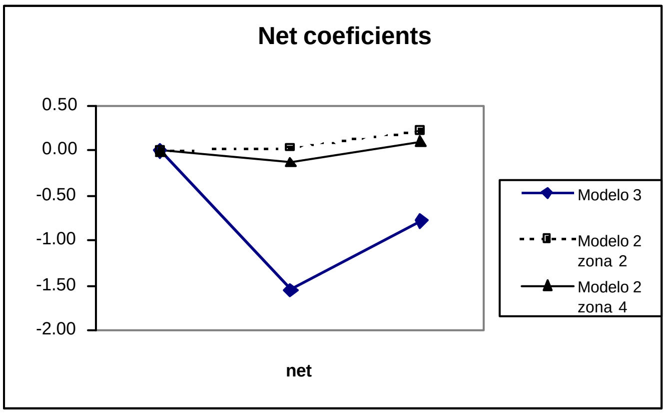	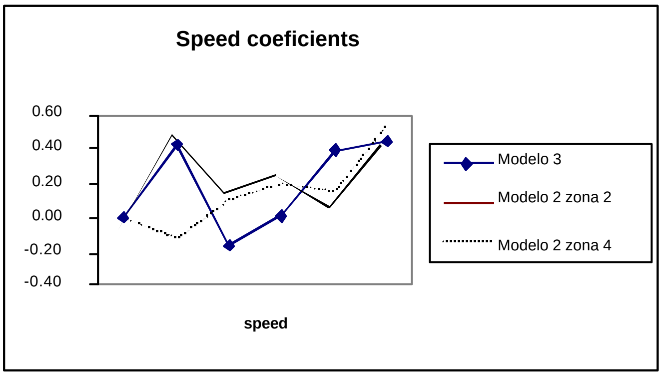
Figure 5: Net coefficients.	*Figure 6: Speed coefficients.*