To be presented in the WKFLAT, Benchmark Workshop on Flatfish Species and Anglerfish, 1-8 March 2012

Modelling Anglefish Discard behaviour in Spanish North Atlantic Coast Fisheries

Juan Santos, Nélida Pérez

Instituto Español de Oceanografía, C.O. Vigo, 36390, Pontevedra, Spain

February 29, 2012

Abstract

Onboard sorting process for anglerfish species in Spanish North Atlantic (ICES VIIIc, IXa) coastal fisheries is reviewed. The length effect analysis across the years sampled reveled an increasing trend in length of first retention (L_{50}) since 2000, the year when Minimum Landing Weight (MLW 500g) were implemented. Specific differences in the length-based sorting process were found, being the less valuable white angler discarded at larger lengths than the black species; further, the analysis found that discard decision is taken at narrower length range for black angler. This results indicate that fishers recognize angler species even at low length sizes, conditioning the degree of adoption of MLW with regards to species relative market value.

1 Introduction

The onboard sorting criteria defining the retained and discarded fractions depends on factors such as legal rules, market value, fishing strategy and targets or hold capacity. Among others, the interaction between market value and legal restriccions defining marketable catch fractions has been recognized as one of the main discarding inducers. Regarding legal restriccions, TAC quota determines what species and how much of these species can be marketed, while Minimum Landing Size/Weigth define the onboard sorting behavior within species; those fishes below the Minimum Landing Size/Weigth are not allowed to be landed and must be thrown back to the sea.

White (*Lophius piscatorius*) and black (*L. budegassa*) anglerfish are high valuable species for fishers operating in ICES Divisions VIIIc and IXa (North Atlantic Iberian waters), being their stock under yearly assessment by the ICES Working Group of Hake, Monk and Megrim (WGHMM). Yearly TACs and Minimum Landing Weight (MLW, which specify a minimun of 500grs for landing individuals [CR 2406/96]) are the most important assessment tools in force for both species. Previous analysis suggest that anglers discards basically depends on market value rather than MLW (Díaz, 2008). The main aim of the present document is to re-analize sorting behaviour over the years sampled by estimating yearly length of first retention (L_{50}) using advanced regression models, and compare the new results with those obtained by Diaz *et al* (2008).

2 Material and Methods

2.1 Data source

Black and white angler catch information from the Spanish North Atlantic (ICES VIIIc, IXa North) coastal fisheries were used to model the onboard sorting behaviour on angler species. Data were obtained by the Spanish Discard Sampling Program -SDSP-, which follows the ICES standards (ICES, 2003; ICES, 2007) for onboard sampling. The sampling protocol are based on a hierarchial design, where the target information, or Ultimate Sampling Unit (USU=Number of individuals sampled) is nested within hauls (Secondary Sampling Unit [SSU]) within trips (Primary Sampling Unit [PSU]). Sampled trips are randomly (or quasirandomly) drawn from predefined stratas (métiers). Once onboard, the observer sistematically choose a number of hauls for sampling, and for every sampled haul, biological information is obtained from both discarded and retained catch fraction. In case of discard data, the observer take a sample (d) from the total discard volume (D), being the biological information in d subsequently raised to haul level by

$$rf_d = \frac{D}{d} \tag{1}$$

2.2 Model settings

Response variable

Discard behavior analysis is carried out by modeling the probability of a given fish to be retained during the onboard sorting process. For this purpose, anglers counts were splitted into `n.r`=Number of retained individuals, and `n.d`=Number of discarded individuals, in order to calculate the proportion of individuals retained within sampled haul within sampled trip;

$$\frac{n.r}{(n.r + n.d)} = \hat{\theta} \tag{2}$$

therefore $\hat{\theta}$ is the empirical proportion of anglerfishes discarded by haul, which only can take values between 0 and 1. $\hat{\theta} = 1$ means that all individuals caugth during the haul j, $j = 1, \cdots J$ were retained onboard, while the opposite $\hat{\theta} = 0$ means that all indiviuals were discarded. In other words, θ is the underlying probability defining retained/discarded fractions;

$$n.r \sim Binomial((n.r + n.d), \theta) \tag{3}$$

Conditional means of θ and their variability are herein estimated under regression methods.

2.3 Estimation method

Generalized Linear Models (McCullagh and Nelder, 1989) relax the Linear Models (LM) assumptions allowing binary variables as model response, such as θ (representing the onboard sorting behavior). In this case, the model structure can be set using the *logit* link function,

$$\theta = \frac{exp(\eta)}{1 + exp(\eta)} \tag{4}$$

with η corresponding to the right-hand side of a classical LM,

$$\eta = \beta_0 + \beta_1 \cdot X_1, \cdots, \beta_p \cdot X_p \tag{5}$$

however, the hierarchial sampling design can implicity produce certain degree of correlation between observations not considered in the GLM framework. This correlation can be given because observations from hauls performed in the same trip could be more similar each others than hauls from different trips. Implicity, this hiphotesys suggest that using a standard binomial GLM could violate the assumption of independence between observations. Generalized Linear Mixed Modelling (GLMM) framework relax the assumption of indepence between observations, allowing correlation structure within data clusters, such as the PSU (trips). Using the GLMM framework makes it possible the joint modeling of:

- Fixed effects: Variables of interest for assessing their effects on the dependent variable.
- Random effects: Clustered sampling units.

2.4 Model Specifications

- model type: Generalized Linear Mixed model (`family=binomial,link=logit`).
- Likelihood function: REML.
- Estimation Method: Laplace aproximation.
- Package: `lme4`.
- Function `glmer()`.
- Response $\frac{n.r}{(n.r+n.d)} = \hat{\theta}$.
- Fixed effects: `length`, `year`.
- Offset=$\frac{rf_r}{rf_d}$
- Random effects: intercept=`trip`; slope= `length`.
- Model selection: `anova(...,test='Chi')` criterion.

3 Results

3.1 White angler models

Models fitted for white angler catch data are given below:

```
rb1<-glmer(cbind(n.r,n.d)~factor(year)+ (1|trip)+offset(log(rf.r/rf.d)),
family=binomial, data=rb)
rb2<-glmer(cbind(n.r,n.d)~factor(year)+ (length|trip)+offset(log(rf.r/rf.d)),
family=binomial, data=rb)
rb3<-glmer(cbind(n.r,n.d)~factor(year)+length+(1|trip)+offset(log(rf.r/rf.d)),
family=binomial, data=rb)
rb4<-glmer(cbind(n.r,n.d)~factor(year)+length+(length|trip)+offset(log(rf.r/rf.d)),
family=binomial, data=rb)
rb5<-glmer(cbind(n.r,n.d)~factor(year)*length+(1|trip)+offset(log(rf.r/rf.d)),
family=binomial, data=rb)
rb6<-glmer(cbind(n.r,n.d)~factor(year)*length+(length|trip)+offset(log(rf.r/rf.d)),
family=binomial, data=rb)
```

The anova table (Table 1) indicates `rb4` as the best model. `rb4` propose that the main retention curve varies in terms of intercept (`trip`) and slope (`length`) across the series (random effects for longitudinal data, see Table 2) . The importance of each of the variance components is given in Table 2 for model `rb4`. The retention ogive intercept (`trip`) is much more variable than the slope (`length`). Both variance components show very high negative correlation value (-0.95), Table 2.

	Df	AIC	BIC	logLik	Chisq	Chi Df	Pr(>Chisq)
rb1	13	2141.49	2236.18	-1057.74			
rb2	15	938.45	1047.71	-454.22	0.00	1	1.0000
rb3	14	536.40	638.38	-254.20	1607.08	1	0.0001
rb4	**16**	**529.40**	**645.94**	**-248.70**	**411.05**	1	**0.0001**
rb5	25	550.94	733.04	-250.47	0.00	9	1.0000
rb6	27	545.12	741.79	-245.56	9.82	2	0.0074

Table 1: Model selection by test (‘Chi’) in Anova table. Best model highlighted in bold format.

Groups	Name	Variance	Std.Dev.	Corr	V6
trip	(Intercept)	63.32	7.95		
	length	0.084	0.29	-0.95	

Table 2: Variance Components and correlations extracted from the best white angler model (`rb4`).

Fixed effects from model `rb4` and their relative significance are given below. Only some years > 2000 significantly differs from the reference 1994:

```
Fixed effects:
                  Estimate Std. Error z value Pr(>|z|)
(Intercept)      -13.77598    1.54841  -8.897  < 2e-16 ***
factor(year)1997   0.03271    2.08737   0.016 0.987498
factor(year)1999  -0.87755   50.83856  -0.017 0.986228
factor(year)2000  -2.16909    1.73416  -1.251 0.211007
factor(year)2003  -2.37912    1.16277  -2.046 0.040749 *
factor(year)2004  -3.74280    1.25439  -2.984 0.002847 **
factor(year)2005  -2.02365    0.86973  -2.327 0.019978 *
factor(year)2006  -2.14947    1.47776  -1.455 0.145795
factor(year)2007  -1.27108    1.02019  -1.246 0.212793
factor(year)2008  -2.81579    1.12345  -2.506 0.012197 *
factor(year)2009  -3.63275    1.08095  -3.361 0.000777 ***
factor(year)2010  -1.58836    1.12148  -1.416 0.156685
length             0.72002    0.05959  12.083  < 2e-16 ***
```

Yearly L_{50} and Selention Range (SR) estimated from `rb4` fixed parameters are given in Table 3. Most of yearly L_{50} estimations since 2000 are significantly higher than the reference L_{50}^{1994}.

	L_{50}	SR
1994	19.13	3.05
1997	19.09	3.05
1999	20.35	3.05
2000	22.15	3.05
2003	22.44*	3.05
2004	24.33*	3.05
2005	21.94*	3.05
2006	22.12	3.05
2007	20.90	3.05
2008	23.04 *	3.05
2009	24.18 *	3.05
2010	21.34	3.05

Table 3: Yearly L_{50} and SR estimated by the best model for white angler (`rb4`)

3.1.1 Black angler models

Models fitted for black angler catch data are given below:

```
rn1<-glmer(cbind(n.r,n.d)~factor(year)+ (1|trip)+offset(log(rf.r/rf.d)),
family=binomial, data=rn)
rn2<-glmer(cbind(n.r,n.d)~factor(year)+ (length|trip)+offset(log(rf.r/rf.d)),
family=binomial, data=rn)
rn3<-glmer(cbind(n.r,n.d)~factor(year)+length+(1|trip)+offset(log(rf.r/rf.d)),
family=binomial, data=rn)
rn4<-glmer(cbind(n.r,n.d)~factor(year)+length+(length|trip)+offset(log(rf.r/rf.d)),
family=binomial, data=rn)
rn5<-glmer(cbind(n.r,n.d)~factor(year)*length+(1|trip)+offset(log(rf.r/rf.d)),
family=binomial, data=rn)
rn6<-glmer(cbind(n.r,n.d)~factor(year)*length+(length|trip)+offset(log(rf.r/rf.d)),
family=binomial, data=rn)
```

The anova table for black angler (Table 4) indicates that `rn4` is the best fitted model. The estructure of this model does not differ from `rb4`. The importance of each of the variance components is given Table 5 for model `rn4` . As in the case of white angler, the retention ogive intercept (`trip`) is much more variable than the slope (`length`). Both variance components also show very high negative correlation value (-0.97).

	Df	AIC	BIC	logLik	Chisq	Chi Df	Pr(>Chisq)
rn1	13	1351.52	1445.93	-662.76			
rn3	14	477.28	578.94	-224.64	876.25	1	0.0001
rn2	15	604.18	713.11	-287.09	0.00	1	1.0000
rn4	16	356.50	472.69	-162.25	249.69	1	0.0001
rn5	25	399.47	581.01	-174.73	0.00	9	1.0000
rn6	27	366.69	562.77	-156.35	36.77	2	0.0001

Table 4: Model selection by Anova test ('Chi')

Groups	Name	Variance	Std.Dev.	Corr	V6
trip	(Intercept)	36.917693	6.07599		
	length	0.063565	0.25212	-0.970	

Table 5: Variance Components and associated s.e. and correlation extracted from the best black angler model (`rn4`).

Fixed effects and their relative significance for model `rn4` are given below. Only some years $\geq$ 2006 significantly differs from the reference 1994:

```
Fixed effects:
                 Estimate Std. Error z value Pr(>|z|)
(Intercept)       -9.8847     1.3144  -7.520 5.46e-14 ***
factor(year)1997   0.8891     0.8222   1.081 0.279491
factor(year)1999  -0.8643     1.6966  -0.509 0.610438
factor(year)2000  -1.2489     0.8317  -1.502 0.133204
factor(year)2003   0.7937     1.8377   0.432 0.665799
factor(year)2004  -0.7338     2.1900  -0.335 0.737558
factor(year)2005  -0.9934     0.9314  -1.067 0.286128
factor(year)2006  -2.1262     0.9093  -2.338 0.019372 *
factor(year)2007  -0.3856     1.1376  -0.339 0.734643
factor(year)2008  -4.1676     1.1075  -3.763 0.000168 ***
factor(year)2009  -1.7213     1.0848  -1.587 0.112572
factor(year)2010  -1.5887     0.7736  -2.053 0.040028 *
length             0.5799     0.0533  10.880  < 2e-16 ***
```

Yearly L_{50} and Selention Range (SR) estimated from `rn4` fixed parameters are given in Table 6. Only three yearly L_{50} estimations since 2006 are significantly higher than the reference L_{50}^{1994}.

	L50	SR
1994	17.05	3.79
1997	15.51	3.79
1999	18.54	3.79
2000	19.20	3.79
2003	15.68	3.79
2004	18.31	3.79
2005	18.76	3.79
2006	20.71*	3.79
2007	17.71	3.79
2008	24.23*	3.79
2009	20.01	3.79
2010	19.78*	3.79

Table 6: Yearly L_{50} and SR estimated by the best model for black angler (`rn4`)

4 Discussion

The analysis carried out by Diaz *et al.* (2008) estimated yearly L_{50} separately for both species; in other words, the authors fitted as many models as years for both species. The GLMM's (`rb4` and `rn4`) herein fitted used all sampled years as data input. White angler model estimate yearly significant differences in the onboard sorting behaviour since 2000, while the black angler model estimate the first significant year in 2006. Both models use 1994 as the reference year. This hypothesis testing is lack in the Diaz *et al.* (2008) approach. Altougth an increase in L_{50} is found for both species since year 2000, this parameter have never reached the value of 31cm, the length associated to 500gr individuals (Diaz *et al.*, 2008), suggesting that species market value is a major factor hindering the full implementation of the MLW regulation. An inter-species L_{50} comparison supports this hyphotesis, given that estimations for the less valuable white angler indicates that this species is generally discarded at larger lengths than black species (see tables 3 and 6). This result suggest that fishers may recognize angler species even at low length sizes, conditioning the degree of MLW adoption with regards to their relative market value.

Further, the analysis found that discard decision is taken at narrower length range (SR) for black angler (tables 3 and 6) . The random effects results indicates that (`trip`) is the main variance component for both species sorting ogives, meaning that the ogive intercepts (varying between trips) are more variable than their shape.

5 References

Díaz, P., Santos, J., Velasco, F., Serrano, A., and Pérez, N. 2008. Anglerfish discard estimates and patterns in Spanish Northeast Atlantic trawl fisheries. – ICES Journal of Marine Science, 65: 1350–1361.

ICES. 2003. Report of the Workshop on Discard Sampling Methodology and Raising Procedures,2 – 4 September 2003, Charlottenlund, Denmark.. Final Report. The ICES Planning Group on Commercial Catch, Discards and Biological Sampling (PGCCDBS).

ICES. 2007. Report of the Workshop on Discard Raising Procedures, 6–9 February 2007, San Sebastian, Spain. ICES CM 2007. ACFM:06. 57 pp.

McCullagh, P., Nelder, J., 1989. *Generalized Linear Models.* Chapman and Hall, London.