

Aplicación de Métodos de Regresión non  
Paramétrica na Estimación de Curvas de  
Selectividade de Procesos Pesqueiros

Juan José Santos Blanco

Este traballo foi presentado por Juan José Santos Blanco, baixo a dirección de Javier Roca Pardiñas, para a consecución do título de Master do Programa Oficial de Postgrado de Estadística e Investigación Operativa, ofertado pola Universidade de Vigo.

É para que así conste, os arriba mencionados firmán baixo esta liña:

Vigo, 25.01.2010

# Resumo

A capacidade selectiva e a eficiencia dos sistemas extractivos determinan en grande medida o impacto ecolóxico e o rendemento económico das actividades pesqueiras. O presente traballo resume os resultados das análises de selectividade e eficiencia pesqueira levadas a cabo sobre casos reais. Os datos empregados foron obtidos en dúas accións piloto, onde se puxeron a proba distintas especificacións pesqueiras. As curvas de regresión empregadas neste traballo foron estimadas baixo técnicas non paramétricas. Os intervalos de confianza e contrastes de hipóteses asociados a cada análise foron construídos utilizando o bootstrap. A caracterización do comportamento pesqueiro das especificacións a estudo permite en última instancia valorar as vantaxes e desvantaxes do seu uso nas actividades de pesca comercial.



# Agradecementos

Me gustaría, en primeiro lugar, agradecer a dispoñibilidade, sacrificio e solidariedade de Javier Roca, o cal afrontou, a piñon grande e plato pequeno, o choque de entendemento entre as linguaxes da bioloxía e a matemática. Gracias tamén ao grupo IBDES por todo o que me ten axudado dende a miña incorporación ás súas filas. Sen a vosa enerxía e humanidade, non podería afrontar con motivación as metas que nos marcamos como grupo. Agradezo especialmente a axuda desinteresada de Hortensia Araujo, sen ela non sería posible rematar a tempo este traballo. Non me podo olvidar de Victor Lopez e Miguel Lustres, Compañeiros de campañas mariñeiras puntuais, e amigos para sempre. Tampouco de Rosa, por traer sempre algunha parrafada dende Canicouva. Por último, agradezo a hospitalidade e entrega das tripulacións dos pesqueiros 'Balayo Portela' e 'Pescaberbes III'. Sen eles non sería posible capturar a información que aquí se analiza.



# Índice general

<b>1. Introducción</b>	<b>9</b>
1.1. Descartes Pesqueiros . . . . .	10
1.2. A Selectividade nos Artes de Pesca . . . . .	12
1.3. Estimación de Procesos Selectivos . . . . .	16
1.4. Obxectivos . . . . .	18
<b>2. Metodoloxía Estatística</b>	<b>21</b>
2.1. Modelos Lineais Xeneralizados(GLM) . . . . .	22
2.1.1. Estimación . . . . .	24
2.2. Modelos Aditivos Xeneralizados (GAM) . . . . .	27
2.2.1. Estimación . . . . .	28
2.2.2. Intervalos de Confianza . . . . .	31
2.2.3. Contrastes de Interaccións Factor-por-Curva . . . . .	34
<b>3. Caso Práctico 1</b>	<b>37</b>
3.1. Antecedentes . . . . .	38
3.2. Acción Piloto . . . . .	39
3.3. Sistema de Pesca . . . . .	39
3.4. Deseño Experimental . . . . .	40
3.5. Obxectivos . . . . .	41
3.6. Análises . . . . .	41
3.6.1. Estimación de Curvas de Retención . . . . .	41
3.6.2. Contrastes de Interacción . . . . .	43
<b>4. Caso Práctico 2</b>	<b>57</b>
4.1. Antecedentes . . . . .	58
4.2. Acción Piloto . . . . .	59
4.3. Sistema de Pesca . . . . .	59

4.4. Diseño Experimental . . . . .	60
4.5. Obxectivos . . . . .	60
4.6. Análises . . . . .	61
4.6.1. Comparación de Capturas . . . . .	61
4.6.2. Contraste de Ratios de Captura . . . . .	61
<b>5. Discusión</b>	<b>69</b>
<b>6. Bibliografía</b>	<b>73</b>



# Capítulo 1

## Introducción

A primeira parte de este apartado inclúe a definición do termino "descartes pesqueiros", expoñendo as causas que os orixinan e os efectos negativos que esta práctica produce na saúde das pesqueiras. Introdúcese de seguido o termo selectividade, revisando brevemente a evolución dos métodos analíticos utilizados para a cuantificación das capacidades selectivas nos sistemas extractivos. Por último, definimos os obxectivos que persegue o presente traballo

## 1.1. Descartes Pesqueiros

Os descartes pesqueiros son definidos pola FAO como “*a materia orgánica de orixe animal capturada, que unha vez izada a bordo, por un motivo ou outro, é devolta ao mar*” (Anónimo, 1996). A presenza desta práctica nas pesqueiras comerciais afecta negativamente ás poboacións de peixes e aos ecosistemas mariños asociados. En termos ecolóxicos, os descartes pesqueiros poden entenderse como o aporte de grandes cantidades de materia orgánica inerte ao ámbito mariño, provocando a redistribución de especies carroñeiras e oportunistas (Olaso et al, 1998; Tasker et al, 2000), e transformando en última instancia a estrutura ecolóxica primixénia (Kaiser e de Groot, 2000), da cal dependen as especies de valor pesqueiro. O incremento da atención nesta actividade levou á comunidade científica a tratar de coñecer o problema dende varias perspectivas, das cales, a estimación da magnitude do volume de descartes nas pesqueiras foi a primeira en interese. Alverson (1994), cuantifica por primeira vez os descartes pesqueiros a nivel global. Para acadar tal obxectivo empregou a información dispoñible de capturas pesqueiras obtidas entre a década dos 80 e os primeiros anos dos 90. Os resultados deste estudio cederon unha estimación de volumen de descarte global comprendido entre os 17.9 e os 39.5 millóns de toneladas. Utilizando métodos de estimación alternativos, Kelleher (2005) actualiza os resultados anteriores. As estimas actuais mostran unha notable diminución ocorrida nos últimos anos, con valores de  $\sim 7.3$  millóns de toneladas / ano. Tal descenso parece motivado tanto pola optimización dos métodos de ponderación de unha cada vez mellor información sobre o problema, como da toma de decisións destinadas a afrontalo (Gabr et al., 2007).

O seguinte paso no esclarecemento e comprensión dos descartes foi a descripción analítica dos factores que inflúen nestes procesos insostibles. Este esforzo investigador desemboca na publicación de artigos de referencia dentro da investigación pesqueira.

Tratando de facilitar a súa comprensión, podemos entender os descartes como o resultado da valoración e posterior toma de decisións que o pescador fai sobre a súa captura. En resumo, ante o limitado espazo existente nas bodegas, as tripulacións tratan de maximizar o potencial económico da captura que é retida para a venda. Baixo esta perspectiva, parte do volume capturado, composto por especies ou clases de tallas carentes de interese nas súas aspiracións comerciais (Rochet e Trenkel, 2003; Catchpole et al.,

2005), son devoltas ao mar. Previa decisión de cal será a parte da captura retida, e cal a descartada, véñense atopando outros factores que condicionan esta práctica:

i. *Regulacións pesqueiras*

Algúns imposicións legais enfocadas á xestión das pesqueiras restrinxen ou anulan a utilización comercial das capturas, favorecendo de maneira indirecta que os descartes pesqueiros prevalezan (Rochet e Trenkel., 2005). Por exemplo, as restricións respecto ás tallas mínimas de desembarco (MLS, *Minimum landing size*), leva ao pescador a descartar aqueles individuos de pequeno tamaño, previamente capturados (Stratoudakis et al., 1998). Por outro lado, cuotas máximas impostas para as especies comerciais (TAC, *Total Allowable Catch*), implican o descarte de aquelas especies con límite de explotación anual superada.

ii. *Disponibilidade de recursos*

Os ecosistemas presentes na área xeográfica onde se establece unha pesqueira determina as especies e as clases de tallas capturadas, condicionando de maneira significativa a cantidade e composición dos descartes (Ordines et al., 2006, Stratoudakis et al., 1998, Trujillo et al., 1997). Do mesmo xeito, a estrutura poboacional das especies mariñas varían a escala temporal, o que incorpora ademais unha compoñente estacional á variabilidade dos descartes (Machias et al., 2001)

iii. *Influencia das técnicas pesqueiras*

Os métodos extractivos utilizados nunha pesqueira determinan o descarte que esta xera (Rochet e Trenkel, 2003), xa que condicionan o tipo e volumen de captura, a partir da cal o pescador selecciona a fracción comercializable. Por exemplo, un mesmo tipo de arte confeccionado con tamaños de malla diferentes, presentará diferencias en canto á composición dos seus descartes, dado que este factor condiciona as cantidades (Howell e Landan, 1987, Murawski, 1996, Gray et al., 2002), diversidade (Murawski, 1996) e composición de tallas de captura (Stratoudakis et al., 2001).

Actualmente, existe unha forte corrente de investigación enfocada á redución dos descartes pesqueiros. Este campo de traballo queda reforzado baixo

a luz dos anteriores esforzos descritivos, que derivaron nunha maior concienciación dos distintos actores implicados no sector pesqueiro. A política, por outro lado, tamén esta a captar a necesidade de afrontar de maneira activa estas actividades insostibles, o que, en última instancia, queda reflexada nunha cada vez menos permisiva lexislación pesqueira.

No 2007, a Comisión europea publicou unha Comunicación sobre futuras políticas de redución de capturas non desexadas [1], co fin de eliminar os descartes nas pesqueiras europeas. Nesta Comunicación, a Comisión anunciou a súa intención de propoñer unha lexislación específica a partires do 2008, ademais de un roteiro cara a súa implementación definitiva.

Un documento non oficial da comisión é publicado no 2008 [2], para convocar a consulta a distintos actores implicados na actividade pesqueira (Estados membros, ONG's, científicos, xestores). O fin da convocatoria era valorar os posibles pasos destinados a eliminar os descartes pesqueiros das augas europeas. As opcións e obxectivos propostos nese documento de consulta foron publicadas en agosto de 2008. Dende ese mesmo ano, dúas pesqueiras europeas foron seleccionadas para acometer a redución dos seus descartes.

Como se tratou con anterioridade, os sistemas de pesca condicionan, tanto cuantitativa, como cualitativamente, os descartes xerados nunha determinada pesqueira. A área de investigación da cal xurde o presente traballo ten como obxectivo a redución dos descartes mediante a mellora da selectividade dos artes de pesca comerciais.

## 1.2. A Selectividade nos Artes de Pesca

De maneira formal, a selectividade pode entenderse como o proceso que causa diferencias entre a composición da captura, e as poboacións que se atopan accesibles ao sistema de pesca (Millar e Fryer 1999; Anónimo, 1996). Tomando como base esa definición, un arte de pesca é selectivo cando minimiza o volume de capturas non desexadas (Massutí et al., 2009; Beutel et al., 2008; Madsen, 2007; Revill et al., 2007; Catchpole et al., 2006; Maartens et al., 2001; Glass 2000). A investigación neste eido trata de xerar sistemas de pesca selectivos alternativos aos sistemas de pesca tradicionais, e en algúns casos os seus produtos chegan a ser adoptados como especificacións obrigada nas pesqueiras comerciais [3]. A implementación de sistemas de pesca

selectivos provocaron en algunhas pesqueiras cambios de tendencia nas series de tempo de descartes, con escenarios temporais previos e póstumos á súa implementación (Enever et al., 2009; Madsen, 2007).

As solucións técnicas destinadas a transformar un arte tradicional en un sistema de alta selectividade poden clasificarse según a estratexia de selección que adoptan (Anónimo, 2004).

- i Selección por tallas: Filtrado mecánico de aqueles individuos que non alcanzan una talla mínima.
- ii Selección activa: Sistemas ideados co fin de separar as especies obxectivo e non obxectivo según diferencias nos seus hábitos ou comportamentos.

O primeiro enfoque trata de evitar as capturas de individuos xuvenís mediante a selección idónea do tamaño de malla co cal se confecciona o arte. Como quedou exposto con anterioridade, o efecto da malla é un factor significativo que condiciona o tipo e volume de descartes, dado que a súa forma, composición e tamaño regulan as posibilidades de escape das especies capturadas (Figura 1.1). En concreto, a relación entre a morfometría do peixe e a forma e tamaño da malla coa cal contacta, condiciona en gran medida a posibilidade de que este sexa filtrado ao exterior (Millar e Fryer, 1999). Nos artes de pesca de arrastre (Figura 1.2), sistema no cal se basea o presente traballo, a selección pasiva ten lugar na parte posterior do conxunto (Fernö e Olsen, 1994), onde se sitúa un saco cego, denominado copo (Figura 1.3), que almacena a captura obtida ao longo do proceso de arrastre. En definitiva, o tamaño de luz da malla co cal está confeccionado o copo, é un factor determinante que condiciona a selectividade do conxunto pesqueiro. Os tamaños mínimos de malla son medidas legalmente impostas ás flotas, sendo utilizada como ferramenta básica na xestión das pesqueiras mundiais (Catchpole et al., 2005; He, 2007).

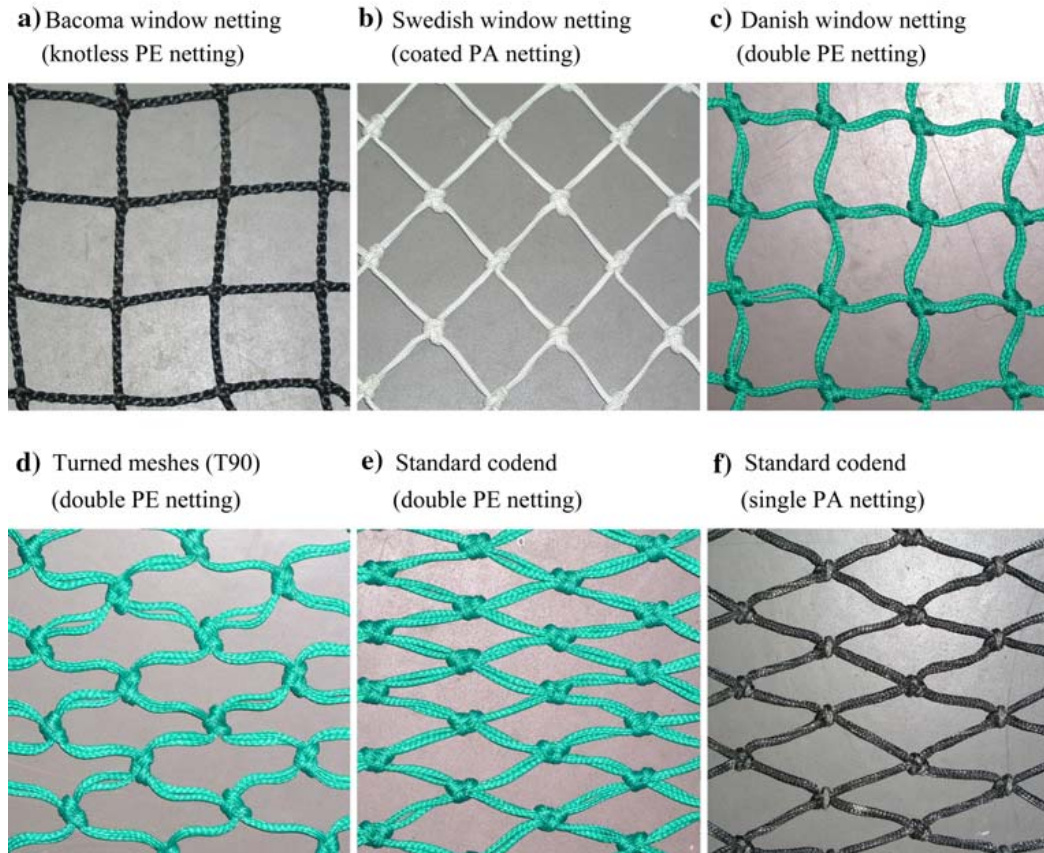


Figura 1.1: Distintos tipos de malla utilizados na pesqueira de bacallao do mar Báltico (Madsen 2007) . A coloración e composición do fio, ademais do tamaño e xeometría da malla condicionan as posibilidades de escape dos seres vivos capturados.

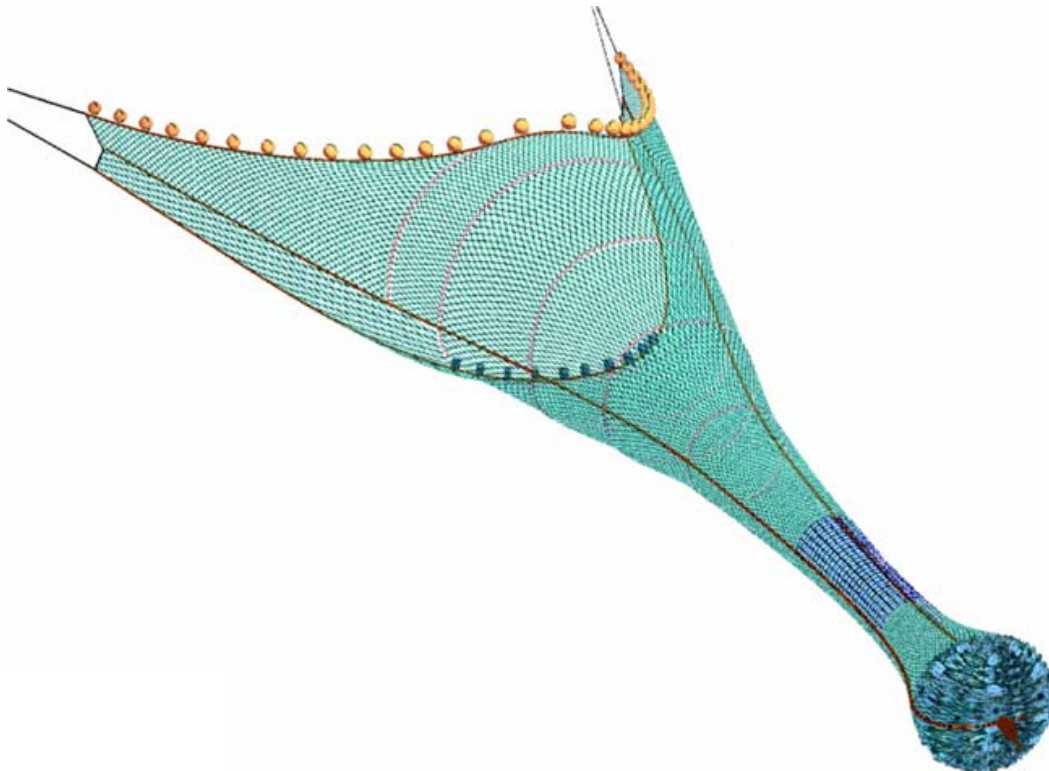


Figura 1.2: Arte de pesca de arrastre. O buque tira polo sistema, que baixo o vector de arrastre, expande a súa estrutura. Da correcta abertura do arte depende a súa eficiencia pesqueira.

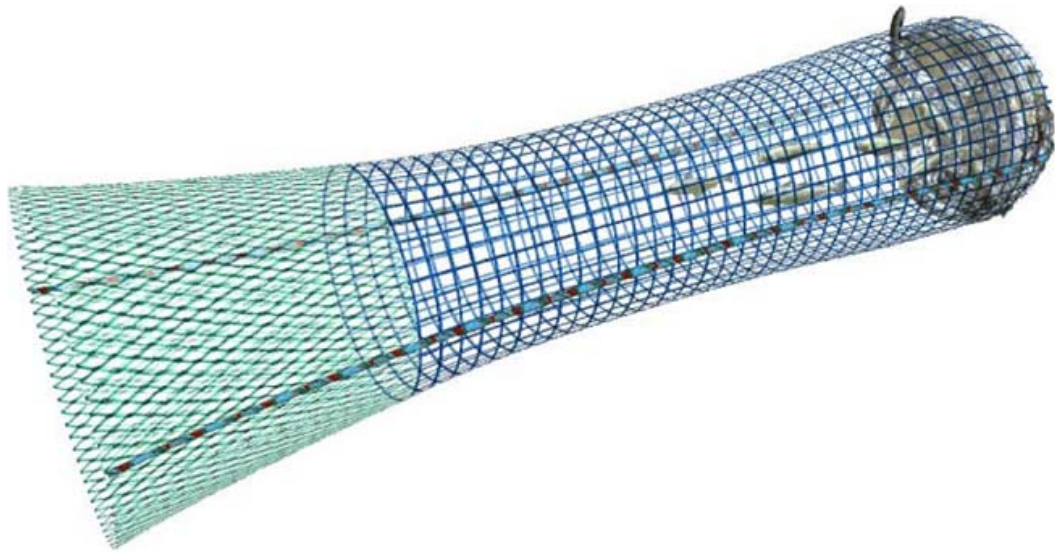


Figura 1.3: Detalle de un copo ensamblado a un arte de arrastre. Na zona posterior do mesmo ten lugar a maior parte da selección pasiva do sistema de pesca.

### 1.3. Estimación de Procesos Selectivos

Retomando o apartado anterior, e evidente a necesidade de avaliar analíticamente o poder selectivo dos sistemas de pesca, na busca das especificacións técnicas óptimas que acaden un compromiso entre a sostibilidade dos procesos pesqueiros, e o seu rendemento económico. Aínda que as primeiras observacións analíticas tiveron lugar a principios do Século XX (Walsh et al.,2002), os avances tecnolóxicos dos anos 40 e 50 para o seguimento dos procedementos pesqueiros levaron a que, en 1955, se creará o ICES (International Council for the Exploration of the Sea) *Committe on comparative fishing*, (denominado a partir de 1983 como Working Group on Fishing Technology and Fish Behaviour, WGFTFB ). Na súa directiva, este WG expón as guías de acción adoptadas:

- Inicio e revisión de investigacións científicas, centradas en tódolos aspectos relacionados co deseño, plantexamento e experimentación de artes de pesca, tanto de aqueles empregados para a estimación da abundancia poboacional, como os destinados a reducir os descartes pesqueiros. Nestas tarefas inclúense a revisión de aspectos etolóxicos



das especies.

- Implicación nas actividades destinadas a desenrolar técnicas de medición e análise das capacidades pesqueiras, tanto dos artes comerciais, como aqueles utilizados na investigación pesqueira, incluído neste punto o deseño de experimentos e a metodoloxía estatística.

Como queda reflexado na directiva, este grupo mantivo un constante interese nos estudos de selectividade, poñendo ademais interese na actualización das distintas metodoloxías estatísticas aplicables, posto que unha mala elección dos procedementos analíticos pode xerar efectos significativos nos resultados dos mesmos (Madsen, 2007).

Revisións metodolóxicas iniciais levados a cabo na década de 1970 (Pope et al., 1975) foron amplamente mellorados en traballos sucesivos nos anos 90 (Fryer, 1991; Millar, 1993; Millar, 1994; Millar e Fryer, 1999), producindo finalmente un marco metodolóxico xeral, que baixo o nome de SELECT (Share Each Lengthclass's Catch Total) facilita aos investigadores pesqueiros, un catálogo variado de deseños experimentais e metodoloxías estatísticas, adaptadas para cada caso. Esta metodoloxía permitiu a globalización da investigación formal dos procesos selectivos en artes de pesca, posibilitando ademais o estudo e recoñecemento de factores que condicionan tales procesos.

A base deste protocolo centrase en tres asuncións:

- A porcentaxe de retención de un arte de pesca é menor no caso de capturas de xuvenís (tallas menores).
- Esta porcentaxe de retención pode variar en función das variantes técnicas de un mesmo arte.
- Algúns artes de pesca poden liberar peixes en un amplo rango de tallas, mentres que outros artes liberan de maneira masiva aqueles individuos por debaixo de tallas concretas.

Unha vez que un peixe contacta co arte de pesca, pasando ao interior do mesmo, xerase unha dicotomía que determina o seu destino final; i) o individuo ben pode quedar retido no fondo do copo, ou ben ii) pode ser liberado a través das mallas. Tales posibilidades poden ser traducidas a un

marco probabilístico, e cando esta probabilidade é graficada en función de un rango determinado de tallas, obtense unha curva de forma sigmoideal, que describe integralmente o poder de retención/selectividade do arte de pesca

En base aos argumentos anteriores, o protocolo SELECT trata de inferir a selectividade mediante a estimación de **curvas de retención**, orixinadas a partires de datos binarios (Anónimo, 1996; Millar e Fryer, 1999)

Outra das aportacións de este protocolo son aquelas técnicas destinadas ao estudio da variabilidade das citadas curvas. As causas de tal variabilidade están conectados a factores non controlados que teñen lugar ao longo dos procesos de pesca (Madsen, 2007; Millar e Fryer, 1999). Ignorar a importancia de esta variabilidade natural pode levar a problemas inferenciais, cando se pretende comparar as capacidades selectivas entre distintos artes. Fryer (1991) desenrolou un método de efectos fixos e aleatorios, co fin de modelar dita variabilidade intrínseca aos procesos pesqueiros. O obxectivo de este deseño é unha estimación realista dos erros estándar asociados aos parámetros de selección. O autor asume que as curvas de selectividade producidas por sucesivos lances (efectos aleatorios) con un mesmo arte (efecto fixo) non é constante, senón que varia de maneira aleatoria arredor de unha curva media. Esta técnica mellora o procedemento clásico de efectos fixos, onde se ignora a variabilidade entre lances, agrupando a información en una sola matriz.

Por outra banda, as técnicas bootstrap (Efron e Tibshirani, 1993) foron utilizadas para o mesmo fin en diversos traballos (Millar, 1993; Munro e Somerton, 2001) sendo incluídas no protocolo SELECT como unha alternativa analítica eficiente.

## 1.4. Obxectivos

Xa exposto o campo de investigación sobre o cal se asenta o presente traballo, enumerase a continuación os obxectivos que este persegue:

- Posta en práctica de metodoloxías analíticas avanzadas na estimación de curvas de selectividade e eficiencia pesqueira, en casos reais.

- Utilización de técnicas inferenciales avanzadas, para a comparación das curvas xeradas por distintas especificacións pesqueiras.



## Capítulo 2

# Metodoloxía Estatística

Este capítulo afronta a revisión dos fundamentos das técnicas estatísticas que serán empregadas nos casos prácticos presentados nos Capítulos 3 e 4. En particular introduciranse os modelos de regresión non paramétricas para resposta binaria, e a metodoloxía bootstrap como una ferramenta para facer inferencia (intervalos de confianza e contrastes de hipótesis) neste tipo de modelos.

## 2.1. Modelos Lineais Xeneralizados (GLM)

O obxectivo de calquera estudo de regresión é atopar un modelo matemático que se axuste aos datos, permitindo unha interpretación razoable da relación entre unha variable resposta  $Y$ , con outro conxunto de  $p$  variables  $\mathbf{X} = (X_1, \dots, X_p)$  (coñecidas como covariables). O interese é polo tanto coñecer a influencia dos valores das covariables  $\mathbf{X}$  na resposta media. Esta cantidade é coñecida como a media condicional, e a denotaremos por  $m(\mathbf{X}) = E[Y|\mathbf{X}]$ .

$$Y = m(\mathbf{X}) + \varepsilon, \quad (2.1)$$

Onde  $\varepsilon$  é definida como a variable de erro. A estrutura elexida para  $m(\mathbf{X})$  determinará o modelo de regresión considerado. Neste senso, os modelos de regresión lineais (LM) supoñen que o efecto da covariable na resposta media é lineal. Explícitamente, a variable resposta  $Y$  é explicada polo vector de covariables  $\mathbf{X}$ , de acordo coa expresión

$$Y = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p + \varepsilon \quad (2.2)$$

sendo  $\beta = (\beta_0, \beta_1, \dots, \beta_p)$  parámetros descoñecidos e  $\varepsilon$  unha variable aleatoria de media cero e varianza  $\sigma^2$ . Dita varianza representa o erro de axuste, ou, dito de outro xeito, a variabilidade de  $Y$  non explicada por  $\mathbf{X}$ .

A estimación dos parámetros  $\beta$  que definen a recta de regresión basease na idea de escoller os estimadores  $\hat{\beta}$  que dean lugar aos residuos máis pequenos. Para tal efecto, e en orde a evitar que se compensen os residuos positivos cos negativos, utilízase a suma cuadrática dos residuos como criterio a minimizar. De tal xeito, dado un conxunto de  $n$  observacións  $(\mathbf{X}_1, Y_1), \dots, (\mathbf{X}_n, Y_n)$  de  $(\mathbf{X}, Y)$ , sendo  $\mathbf{X}_i = (X_{i1}, \dots, X_{ip})$ , a estimación mínimo-cuadrática da recta de regresión consiste en buscar os valores  $\beta = (\beta_0, \beta_1, \dots, \beta_p)$  que minimizan a suma de residuos ao cadrado:

$$\sum_{i=1}^n (Y_i - (\beta_0 + \beta_1 X_{i1} + \dots, \beta_p X_{ip}))^2 \quad (2.3)$$

Os LM son modelos de unha gran aplicación histórica, debido á súa sinxela concepción e utilización, ademais da intuitiva interpretación dos seus resultados. O efecto de cada covariable  $X_j$  ben dado por un único coeficiente  $\beta_j$ ). Sen embargo, a utilización dos LM non está exento de inconvenientes. Este tipo de modelos están pensados para ser empregados sobre variables de

natureza continua, e ademais asumese a priori que a variable  $\varepsilon$  se distribue de acordo con unha normal de media cero e varianza constante, que non depende do valor das covariables  $\mathbf{X}$ . Claramente, tales asuncións limitan a correcta especificación deste tipo de modelos a respostas gaussianas con erros homocedásticos. Pero cando se afrontan análisis de datos reais, e habitual que se produzan violacións nalgunha (ou todas) de estas asuncións, e, polo tanto, os resultados obtidos con estes modelos poden ser erróneos.

Un caso particular deste fenómeno acontece en aqueles estudos onde a variable resposta non pode máis que tomar un valor dicotómico, indicando a ocorrencia ( $Y_i = 1$ ), ou non ocorrencia ( $Y_i = 0$ ) de un suceso de interese. Neste caso, a media da resposta coincide coa probabilidade de que a resposta  $Y$  se manifieste no valor 1, é decir,  $E[Y|\mathbf{X}] = p(Y = 1|\mathbf{X})$ , que en diante será denotada simplemente por  $p(\mathbf{X})$ . A utilización dos LM clásicos nesta situación,

$$Y = p(\mathbf{X}) + \varepsilon \quad (2.4)$$

viola as asuncións de :

1. normalidade: O erro  $\varepsilon$  é neste caso unha variable discreta que sólo pode tomar dous posibles valores: Se  $Y = 1$  entón  $\varepsilon = 1 - p(\mathbf{X})$  con probabilidade  $p(\mathbf{X})$ , e se  $Y = 0$ , entón  $\varepsilon = -p(\mathbf{X})$  con probabilidade  $1 - p(\mathbf{X})$ .
2. homocedasticidade:  $\varepsilon$  é unha variable de media cero e varianza  $Var[Y|\mathbf{X}] = p(\mathbf{X})(1 - p(\mathbf{X}))$ , que depende do vector de covariables  $\mathbf{X}$ .
3. Ademais, a utilización dun modelo clásico para respostas binarias pode provocar a obtención de resultados absurdos, xa que a resposta  $Y$  solo pode tomar dous posibles valores, cero ou un, e esta restricción non é imposta nos modelos clásicos, nos cales  $p(\mathbf{X})$  pode tomar calqueira valor na recta real  $\mathfrak{R}$ , e polo tanto non verificar a restricción  $0 \leq p(\mathbf{X}) \leq 1$ .

Unha posible forma de evitar estas limitacións ben dada por transformación dos datos (Cox, 1970), que permitan obter unha distribución normal con varianza constante. Sen embargo, non sempre é facil a obtención de dita transformación, sendo este un problema aínda aberto.

Máis recentemente aparecen os modelos de regresión xeneralizados, como extensións matemáticas dos modelos clásicos, que non forzan aos datos a escalas artificiais, e que permiten estruturas para a varianza que dependen dos valores das covariables. Este tipo de modelos superan as limitacións anteriores adaptando a resposta, de modo que se adecue á estrutura do modelo empregado. No caso da resposta binaria, a transformación de  $p(\mathbf{X})$  dada por

$$Odds = \frac{p(\mathbf{X})}{1 - p(\mathbf{X})}, \quad (2.5)$$

coñecida como *odds ratio*, pode tomar valores no intervalo  $[0, +\infty)$ . Se ademais levamos a cabo unha transformación logarítmica,

$$\log Odds(\mathbf{X}) = \log \frac{p(\mathbf{X})}{1 - p(\mathbf{X})} \quad (2.6)$$

o rango de posibles valores ampliase a  $(-\infty, +\infty)$ . Polo tanto, a transformación  $\log Odds(\mathbf{X})$ , coñecida como transformación *logit*, pode ser axustada mediante unha ecuación de regresión tradicional,

$$\log Odds(\mathbf{X}) = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p \quad (2.7)$$

ou de maneira equivalente,

$$p(\mathbf{X}) = \frac{\exp(\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p)}{1 + \exp(\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p)} \quad (2.8)$$

dando lugar ao coñecido como modelo loxístico.

### 2.1.1. Estimación

Sexa  $(\mathbf{X}_1, Y_1) \dots (\mathbf{X}_n, Y_n)$  unha mostra de observacións, con vector de covariables  $\mathbf{X}_i = X_{i1}, \dots, X_{ip}$ , seguindo o modelo loxístico

$$p(Y = 1|\mathbf{X}_i) = p_i(\beta) = \frac{\exp(\beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip})}{1 + \exp(\beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip})}$$

Para a estimación dos coeficientes  $\beta = \beta_0, \dots, \beta_p$  emplease o método de máxima verosimilitude. Trátase de obter os coeficientes  $\beta$  que maximizan o logaritmo da función de verosimilitude:



$$l(\beta_0, \dots, \beta_p) = \sum_{i=1}^n Y_i \log p_i(\beta) + (1 - Y_i) \log (1 - p_i(\beta)) \quad (2.9)$$

As ecuacións resultantes do problema de maximización anterior son non lineais, sendo necesario utilizar algún método iterativo. Estes métodos consisten en dar unha solución inicial  $\hat{\beta}^0$  de  $\beta$  e ir dando iterativamente estimacións  $\hat{\beta}^0, \hat{\beta}^1, \hat{\beta}^2 \dots$  que deben converxer a un valor  $\hat{\beta}$  denominado estimador de máxima verosimilitud de  $\beta$ .

Un dos métodos máis utilizado é o algoritmo "Fisher scoring" (McCullagh e Nelder, 1989) que pode ser visto, en cada iteración, como un problema de mínimos cadrados ponderados. Os pasos de este algoritmo son os que seguen:

- **Inicio:** Calcular os coeficientes iniciais  $\hat{\beta}_0^0 = \log \frac{\bar{Y}}{1-\bar{Y}}, \hat{\beta}_1^0 = 0, \dots, \hat{\beta}_p^0 = 0$  sendo  $\bar{Y} = n^{-1} \sum_{i=1}^n Y_i$ .
- **Paso 1.** Calcular, para  $i=1, \dots, n$ , as respostas linealizadas

$$Z_i = \eta_i(\hat{\beta}^0) + \frac{Y_i - p_i(\hat{\beta}^0)}{p_i(\hat{\beta}^0) (1 - p_i(\hat{\beta}^0))}$$

e os pesos

$$W_i = \frac{1}{p_i(\hat{\beta}^0) (1 - p_i(\hat{\beta}^0))}$$

sendo

$$\eta_i(\hat{\beta}^0) = \beta_0 + \hat{\beta}_1^0 X_{i1} + \dots + \hat{\beta}_p^0 X_{ip}$$

e

$$p_i(\hat{\beta}^0) = \frac{\exp(\eta_i(\hat{\beta}^0))}{1 + \exp(\eta_i(\hat{\beta}^0))}$$

- **Paso 2.** Obter (axustando un modelo lineal ponderado) os coeficientes actualizados

$$\hat{\beta} = (\hat{\beta}_0 \dots \hat{\beta}_p)^t = (\mathbf{X}_\bullet^t \mathbf{W}_\bullet \mathbf{X}_\bullet)^{-1} \mathbf{X}_\bullet^t \mathbf{W}_\bullet \mathbf{Z}$$

con

$$\mathbf{X}_\bullet = \begin{pmatrix} 1 & X_{11} & \dots & X_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & \dots & X_{np} \end{pmatrix} \quad \text{e} \quad \mathbf{W}_\bullet = \text{diag}(W_1, \dots, W_n)$$

- **Paso 3.** Repetir o **Paso 1** e **Paso 2** reemplazando as estimacións iniciais,  $\hat{\beta}^0$ , polas finais,  $\hat{\beta}$ , ata acadar a converxencia.

Na converxencia do algoritmo anterior obtense a seguinte aproximación gaussiana

$$\begin{pmatrix} \hat{\alpha}_0 \\ \hat{\alpha}_1 \\ \vdots \\ \hat{\alpha}_p \end{pmatrix} \approx N \left( \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_p \end{pmatrix}, (\mathbf{X}_\bullet^t \mathbf{W}_\bullet \mathbf{X}_\bullet)^{-1} \right)$$

Este resultado permite construír intervalos de confianza e realizar contrastes de hipótesis sobre os coeficientes do modelo. Por exemplo:

- O IC 95 % para  $\alpha_j$  ven dado por

$$\left( \hat{\alpha}_j \mp 1,96 \hat{\Sigma}_j \right)$$

sendo  $\Sigma_j$  o elemento j-ésimo da diagonal principal de  $\Sigma$

- Para contrastar a hipótese nula  $H_0 : \alpha_j = 0$  utilizamos o estatístico de contraste

$$D = \frac{\hat{\alpha}_j}{\hat{\Sigma}_j} \approx N(0, 1) \text{ baixo } H_0$$

e o p-valor do contraste será

$$p\text{-valor} = 2 \cdot (N(0, 1) \geq |d|)$$

## 2.2. Modelos Aditivos Xeneralizados (GAM)

Como acabamos de ver, os GLM son unha extensión dos modelos lineais que cobren o estudio de diversos tipos de datos, como por exemplo os datos binarios. Sen embargo, en ocasións, as modelizacións obtidas mediante técnicas clásicas paramétricas resultan demasiado ríxidas -ou incluso inadecuadas- para o problema de interés. Ante este tipo de situacións, é necesario o desenvolvemento e aplicación de modelos máis xerais e flexibles, que permitan unha correcta modelización matemática.

Nos últimos anos xurde unha liña de investigación no campo da estadística funcional non paramétrica que permite a aplicación de modelos máis xerais. Hastie e Tibshirani (1990) propoñen, evitando a suposición de linealidade, a utilización dos modelos aditivos xeneralizados (GAM). Os GAM poden ser vistos como a evolución natural dos GLM, dado que aportan un novo contexto de flexibilidade, ampliando notablemente a súa aplicación en casos prácticos.

Nos GLM supoñíase un efecto lineal de cada covariable  $X_j$  na resposta  $Y$  (en escala loxística). Esta suposición pode restrinxir a efectividade deste tipo de modelos na descrición e predicción de sucesos, nos casos onde os efectos das covariables non necesariamente seguen ese tipo de asociación. Os modelos GAM reconsideran a compoñente sistemática  $\eta = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p$  permitindo unha estrutura non lineal, introducindo funcións descoñecidas (non paramétricas) en lugar dos clásicos efectos lineais. Formalmente, os GAM loxísticos veñen dados pola expresión

$$p(\mathbf{X}) = \frac{\exp(\beta_0 + f_1(X_1) + \dots + f_p(X_p))}{1 + \exp(\beta_0 + f_1(X_1) + \dots + f_p(X_p))} \quad (2.10)$$

Onde se pode observar como as constantes asociadas as covariables  $\mathbf{X}$  son suplantadas por funcións  $f_1, \dots, f_p$  sen especificar, e polo tanto, descoñecidas. En resumo, os GAM estenden aos GLM incorporando:

- A **flexibilidade** da regresión non paramétrica.
  - O efecto de cada covariable non segue unha forma paramétrica fixa, senon que depende de unha función totalmente descoñecida á cal unicamente se lle esixe un certo grado de suavidade.
- a **interpretabilidade** dos modelos paramétricos.

- A suposición de aditividade fai que o efecto de cada covariable veña dada por unha única curva.

Tales avances fan dos GAM unha das ferramentas con maior capacidade para a análise de datos en distintos campos de investigación. Esta é a razón pola cal son seleccionadas no presente traballo de investigación pesqueira.

### 2.2.1. Estimación

Para a estimación dos GAM é habitual utilizar unha modificación do algoritmo Fisher Scoring que vimos de expoñer na sección 2.1.1. (Estimación dos GLM). A adaptación ben dada no **Paso 2**, que, en lugar de axustar un modelo lineal ponderado, axustará un modelo aditivo ponderado. Para a estimación das funcións parciais que compoñen dito modelo aditivo, será necesario empregar algún método de suavización. Entre outros, pódense destacar: suavizadores kernel (Nadaraya, 1964; Watson, 1964), suavizadores polinómicos locais, Cubic smoothing splines (Wahba, 1990; Hastie e Tibshirani, 1990) ou splines de regresión penalizados (Ruppert et al., 2003; Wood, 2006).

A partir da mostra  $\{\mathbf{X}_i, Y_i\}_{i=1}^n$  a estimación do GAM en 2.10 obtense de acordo ao seguinte procedemento iterativo:

- **Inicio:** Calcular as estimacións iniciais

$$\hat{\beta}_0^0 = \log \frac{\bar{Y}}{1 - \bar{Y}}, \hat{f}_1^0 = 0 \dots, \hat{f}_p^0 = 0$$

- **Paso 1:** Obter as respostas linealizadas e os pesos

$$Z_i = \hat{\eta}_i^0 + \frac{Y_i - \hat{p}_i}{\hat{p}_i(1 - \hat{p}_i)} \quad \text{e} \quad W_i = \frac{1}{\hat{p}_i(1 - \hat{p}_i)}$$

a partires das estimacións iniciais

$$\hat{\eta}_i^0 = \hat{\beta}_0^0 + \hat{f}_1^0(X_{i1}) + \dots + \hat{f}_p^0(X_{ip})$$

e

$$\hat{p}_i = \frac{\exp(\eta_i^0)}{1 + \exp(\eta_i^0)}$$

- **Paso 2:** Axustar un modelo aditivo de  $Z$  sobre  $\mathbf{X}$ , ponderado por  $W$ , obtenndo para  $i = 1, \dots, n$  as actualizacións

$$\hat{f}_1(X_{i1}), \dots, \hat{f}_p(X_{ip})$$

e

$$\hat{\eta}_i = \hat{\alpha} + \hat{f}_1(X_{i1}) + \dots + \hat{f}_p(X_{ip})$$

- **Paso 3:** Repetir os **Pasos 1 e 2** reemplazando  $\hat{\eta}_i^0$  por  $\hat{\eta}_i$  hasta a converxencia.

O procedemento de estimación do modelo aditivo no **Paso 2** depende do número e da natureza das funcións parciais incluídas no modelo: Cando as funcións parciais son paramétricas, p.ex.  $f_j(X_j) = \alpha_j X_j$ , o modelo de regresión aditivo resolverse usando mínimos cadrados ponderados, e o GAM é equivalente ao GLM. En cambio, cando ditas funcións non son paramétricas, as actualizacións  $\hat{f}_j$  son obtidas axustando un modelo aditivo ponderado.

### P-splines

Neste traballo, para a estimación do modelo aditivo no **Paso 2**, consideraremos o marco metodolóxico da Regresión Spline Penalizada desenrolado por Wood (2006), que dan lugar á seguinte estrutura:

$$f_j(x_j) = \sum_{k=1}^{K_j} b_{jk}(x_j) \beta_{jk} \quad (2.11)$$

onde

- $b_{jk}(\cdot)$ ,  $k = 1, \dots, K_j$  son funcións coñecidas,
- $K_j$  indica o número de nodos, e
- $\beta_{jk}(\cdot)$ ,  $k = 1, \dots, K_j$  son parámetros descoñecidos.

Con esta representación, cada función  $f_j$  pasa a ter unha forma paramétrica dependendo unicamentedos dos parámetros  $\beta_{jk}(\cdot)$ ,  $k = 1, \dots, K_j$ . Así, o modelo aditivo no **Paso 2** pode ser visto como un modelo lineal

$$\eta_i = \sum_{k=1}^{K_1} b_{1k}(x_{1i})\beta_{1k} + \dots + \sum_{k=1}^{K_p} b_{pk}(x_{pi})\beta_{pk} = \mathbf{X}_i\beta \quad (2.12)$$

Pódese pensar, por lo tanto, en estimar o modelo mediante as técnicas de estimación dos GLM. Sen embargo, sería necesario determinar previamente o número de nodos e a posición dos mesmos, para cada una de las covariables, o que non sempre e sinxelo.

Unha forma de evitar estos problemas consiste en seleccionar un número de nodos suficientemente alto, e controlar o grado de suavización (evitando así o sobre-axuste), incorporando unha penalización no proceso de estimación dos coeficientes de (2.12). Os coeficientes  $\beta$  son estimados minimizando a seguinte suma de residuos cadrados penalizados

$$\sum_{i=1}^n W_i (Z_i - \mathbf{X}_i\beta)^2 + \lambda_1 \int (f_1''(X_1))^2 dx_1 + \dots + \lambda_p \int (f_p''(X_p))^2 dx_p \quad (2.13)$$

onde a integral da derivada segunda ao cadrado mide o grado de curvatura de cada unha das funcións suaves, penalizada polos valores de  $\lambda_j$  (parámetros de suavizado).

- Valores de  $\lambda_j \rightarrow \infty$  levan a estimar  $f_j$  como unha recta,
- mentras que valores de  $\lambda_j=0$  daría lugar a unha estimación non penalizada e dará lugar á interpolación dos datos.

A determinación de  $\lambda_j$  convertese, polo tanto, no proceso máis determinante do método de regresión spline penalizada. Existen para tal efecto distintos criterios para a elección óptima de estos parámetros. Caben destacar aqueles baseados na minimización do erro de predicción, entre eles, a Validación Cruzada (CV), Validación Cruzada Xeralizada, Criterio de infomarción de Akaike (AIC), e o criterio de información Baiesiano (BIC). En caso de modelos baseados en respostas non gaussianas, como as binarias ou de Poisson, existe a alternativa denominada como criterio UBRE (Un-Biased Risk Estimator).

Os resultados dos casos prácticos que se expoñerán no presente traballo foron obtidos utilizando o paquete `mgcv` (Wood, 2000) dispoñible na librería do software libre R (R Development Core Team 2009).

### 2.2.2. Intervalos de Confianza

A estimación de  $\hat{f}_j$  non é suficiente para determinar o efecto que describe. Para cada punto da mostra  $X_{ij}$  interesa coñecer o intervalo onde se sitúe o hipotético valor  $f_j(X_{ij})$  con unha determinada probabilidade. É dicir, tratase de calcular os límites  $f_j^{\min}(X_{ij})$  y  $f_j^{\sup}(X_{ij})$  de modo que se cumpra a seguinte probabilidade

$$p(f_j^{\min}(X_{ij}) \leq f_j(X_{ij}) \leq f_j^{\sup}(X_{ij})) = 1 - \alpha$$

determinándose entón

$$[f_j^{\min}(X_{ij}), f_j^{\sup}(X_{ij})],$$

como un intervalo de confianza puntual para  $f_j(X_{ij})$  con nivel de confianza de  $1 - \alpha$ .

Para a construción de ditos intervalos, é necesario coñecer a distribución das estimacións  $\hat{f}_j$ . Sen embargo, o algoritmo de estimación local scoring, utiliza técnicas de suavización non paramétrica, ademais, a súa natureza iterativa complica notablemente a obtención de resultados asintóticos.

A rutina `gam` do paquete `mgcv` permite obter os erros estándar aproximados  $\hat{\sigma}_j(X_{ij})$  das estimacións  $\hat{f}_j(X_{ij})$ . Asumindo normalidade, obtendose

$$\left( \hat{f}_j(X_{ij}) - z_{1-\alpha/2} \hat{\sigma}_j(X_{ij}), \hat{f}_j(X_{ij}) + z_{1-\alpha/2} \hat{\sigma}_j(X_{ij}) \right),$$

un intervalo de confianza de nivel  $1 - \alpha$  para  $f_j(X_{ij})$ , sendo  $z_p$  o cuantil de orden  $p$  da distribución normal estandarizada.

Sen embargo, hasta a actualidade non existe unha teoría cerrada que permita facer inferencia neste contexto, e a utilización das técnicas de remostraxe bootstrap foron suxeridas ( p.ex. Härdle et al. (2004)) como unha boa alternativa.

### Técnicas Bootstrap

O incremento do poder computacional ocorrido nos últimos anos trouxo consigo unha maior consideración aos métodos de simulación virtual, a partir dos cales nace o bootstrap (Efron,1979). Esta técnica fundamentase no

tratamento de un alto número de remostras virtuais, xeradas a partir da colección de datos empíricos. A súa potencia é funcionabilidade queda demostrada na avaliación da eficiencia de un estimador, ou a construción de Intervalos de Confianza, test de hipóteses, etc... (Efron e Tibshirani, 1993). O contexto bootstrap mostrase polo tanto como unha alternativa consistente as técnicas clásicas, que, apoiadas no Teorema Central do Límite, están fortemente condicionadas á aleatoriedade e representatividade da mostra respecto á poboación. Ademais, os métodos bootstrap simplifican enormemente o cálculo de estimadores caracterizados por regras asintóticas dificilmente recoñecibles.

No contexto de regresión que nos ocupa, existen distintos procedementos para a xeración das réplicas bootstrap. Así, para modelos de regresión de resposta continua, é habitual utilizar o wild bootstrap (Härdle e Mammen, 1993; Härdle e Marron, 1991). Este tipo de remostraxe permite heterocedasticidade na varianza dos residuos. Dado un modelo de regresión de resposta continua

$$Y_i = m(\mathbf{X}_i) + \epsilon_i$$

O procedemento wild bootstrap para a construción de intervalos de confianza para  $m(\mathbf{X})$  consiste nos seguintes pasos:

- **Paso 1.** Para  $i = 1, \dots, n$  calculanse os residuos

$$\hat{\epsilon}_i = Y_i - \hat{m}(\mathbf{X}_i)$$

- **Paso 2.** Para  $b = 1, \dots, B$  (p.ej.  $B = 1000$ ) obtéñense as estimacións  $\hat{m}^{*b}(\mathbf{X}_i)$  a partires da mostra  $\{\mathbf{X}_i, Y_i^{*b}\}_{i=1}^n$  con respostas bootstrap  $Y_i^{*b}$  xeradas en base a

$$Y_i^{*b} = \hat{m}(\mathbf{X}_i) + \hat{\epsilon}_i^{*b} \quad i = 1, \dots, n$$

onde  $\hat{\epsilon}_i^*$  obtense a partires da seguinte distribución de probabilidade

$$\hat{\epsilon}_i^{*b} = \begin{cases} \frac{\hat{\epsilon}_i(1-\sqrt{5})}{2} & \text{con probabilidade } \frac{5+\sqrt{5}}{10} \\ \frac{\hat{\epsilon}_i(1+\sqrt{5})}{2} & \text{con probabilidade } \frac{5-\sqrt{5}}{10} \end{cases}$$

Nótese que cada residuo é xerado a través dunha distribución que trata de imitar a distribución de  $\hat{\epsilon}_i$  verificando  $E(\hat{\epsilon}_i^*) = 0$ ,  $E(\hat{\epsilon}_i^{*2}) = \hat{\epsilon}_i^2$ , e  $E(\hat{\epsilon}_i^{*3}) = \hat{\epsilon}_i^3$ .



- Finalmente, o intervalo de confianza para  $m(\mathbf{X})$  con nivel de confianza  $1 - \alpha$  ven dado por

$$(\hat{m}^{\alpha/2}(\mathbf{X}), \hat{m}^{1-\alpha/2}(\mathbf{X}))$$

sendo  $\hat{m}^p$  o cuantil de orden  $p$  das estimacións  $\hat{m}^{*1}(\mathbf{X}), \dots, \hat{m}^{*B}(\mathbf{X})$ .

### Bootstrap Binario

O método de remostraxe wild bootstrap foi deseñado sobre todo para modelos de regresión de resposta continua. Pero esta metodoloxía non se adapta a outro tipo de respostas. Por exemplo, en modelos de regresión de resposta binaria, obteriamos respostas bootstrap  $Y_i^*$  distintas de cero o a un, como cabería esperar. Debido a estas consideracións, desenvolveronse novos métodos de remostraxe deseñados para distintos tipos de resposta. En particular, neste traballo utilizarase o chamado bootstrap binario, que como o seu nome indica, está deseñado para modelos de regresión de resposta binaria.

A continuación esquematizanse os pasos habituais na construción de intervalos de confianza, baseados na aplicación de bootstrap binario sobre os GAM de resposta binaria. Considéramos o GAM loxístico dado en 2.10 do cal interesa algunha cantidade

$$R = R(\alpha, f_1, \dots, f_p)$$

como por exemplo  $R = f_j(x)$  ou  $R = p(\mathbf{X})$ . O IC bootstrap para devandita cantidade  $R$  obtense de acordo ao seguinte procedemento bootstrap:

- **Paso 1:** A partires da mostra orixinal  $\{\mathbf{X}_i, Y_i\}_{i=1}^n$  obtense a estimación de  $R$

$$\hat{R} = R(\hat{\alpha}, \hat{f}_1, \dots, \hat{f}_p)$$

e as estimacións piloto das medias condicionais

$$\hat{p}(\mathbf{X}_1), \dots, \hat{p}(\mathbf{X}_n)$$

- **Paso 2:** Para  $b = 1, \dots, B$ , xerase a mostra bootstrap  $\{\mathbf{X}_i, Y_i^{*b}\}_{i=1}^n$  con

$$Y_i^{*b} \sim \text{Bernoulli}(\hat{p}(\mathbf{X}_i)),$$

e se calcula a correspondente estimación de  $\hat{R}^{*b}$ .

- **Paso 3:** Unha vez finalizadas as repeticións bootstrap, o intervalo de confianza para  $R$  con nivel de confianza  $1 - \alpha$  ven dado por

$$\left( \hat{R}^{\alpha/2}, \hat{R}^{1-\alpha/2} \right)$$

sendo  $\hat{R}^p$  o cuantil de orden  $p$  dos valores  $\hat{R}^{*1}, \dots, \hat{R}^{*B}$

### 2.2.3. Contrastes de Interaccións Factor-por-Curva

En moitas situacións, o efecto de cada covariable continua  $X_j$  na respostr  $Y$  pode variar cos distintos niveis  $\{1, \dots, M\}$  de un factor  $Z$ . Neste tipo de situacións, o GAM logístico con interaccións factor-por-curva ben dado por

$$p(Z, \mathbf{X}) = \frac{\exp \left\{ \alpha + \sum_{j=1}^p f_j(Z, X_j) \right\}}{1 + \exp \left\{ \alpha + \sum_{j=1}^p f_j(Z, X_j) \right\}}$$

con

$$f_j(Z, X_j) = \underbrace{f_j^0(X_j)}_{\text{efecto principal}} + \overbrace{\begin{cases} f_j^1(X_j) & \text{se } Z = 1 \\ \vdots \\ f_j^M(X_j) & \text{se } Z = M \end{cases}}^{\text{interacción}}$$

Neste contexto será de interés dispoñer de contrastes estatísticos destinados a verificar se o efecto de unha covariable continua  $X_j$  depende ou non dos niveis do factor  $Z$ . É dicir, asumindo o modelo anterior, a hipótese nula é

$$\mathbf{H}_0 : f_j(X_j) = f_j^0(X_j)$$

frente á alternativa

$$\mathbf{H}_1 : f_j(Z, X_j) = f_j^0(X_j) + \begin{cases} f_j^1(X_j) & \text{si } Z = 1 \\ \vdots \\ f_j^M(X_j) & \text{si } Z = M \end{cases}$$

Con ese obxectivo, a partires de unha mostra  $\{Z_i, \mathbf{X}_i, Y_i\}_{i=1}^n$  podemos considerar o estadístico de contraste,

$$T = \sum_{i=1}^n \left( \sum_{l=1}^M \hat{f}_j^l(X_{ij}) I_{\{Z_i=l\}} \right)^2$$

baseado na estimación "directa" dos efectos asociados aos niveles do factor  $Z$ . O contraste rexeitará a hipótese nula para valores de  $T$  suficientemente grandes. En particular, para un nivel de significación  $1 - \alpha$ , rexeitase  $H_0$  se  $T \geq T^\alpha$  donde  $T^\alpha$  é o cuantil  $1 - \alpha$ , da distribución de  $T$  baixo  $H_0$ .

Sen embargo, na práctica, é difícil obter a distribución asintótica do estadístico de contraste  $T$ . Non é posible, polo tanto, o cálculo directo dos cuantiles  $T^\alpha$  necesarios para a implementación do test de interacción. É aquí onde o bootstrap volta a ser unha ferramenta práctica para afrontar estes problemas asociados á estadística clásica.

Os pasos de este procedemento son os que seguen:

- **Paso 1:** A partires dos datos  $\{Z_i, \mathbf{X}_i, Y_i\}_{i=1}^n$  obtemos o valor de  $T_k$  e as estimacións  $\hat{p}_0(Z_1, \mathbf{X}_1), \dots, \hat{p}_0(Z_n, \mathbf{X}_n)$  baixo  $\mathbf{H}_0$  das probabilidades  $p(Z_1, \mathbf{X}_1), \dots, p(Z_n, \mathbf{X}_n)$ .
- **Paso 2:** Para  $b = 1, \dots, B$  xeramos a mostra bootstrap

$$\{Z_i, X_i, Y_i^{*b}\}_{i=1}^n \text{ con } Y_i^{*b} \sim \text{Bernoulli}(\hat{p}_0(Z_i, \mathbf{X}_i)),$$

Obtendo o valor  $T_k^{*b}$  do estadístico de contraste.

- **Paso 3:** Finalmente, rexeitamos a hipótese nula se  $T > T^{*(\alpha)}$  onde  $T^{*(\alpha)}$  é o cuantil empírico de orde  $1 - \alpha$  dos  $B$  valores  $T^{*1}, \dots, T^{*B}$  obtidos no **Paso 2**.

O test directo mostrado nesta sección, combinado co bootstrap, son as técnicas empregadas (a efectos de inferencia) en torno aos contrastes factor por curva, especificados nos casos prácticos que se expoñen nos seguintes capítulos da presente memoria.



# Capítulo 3

## Caso Práctico 1

### 3.1. Antecedentes

Os resultados dos procesos pesqueiros dependen de distintos factores que condicionan as capturas de un arte. Entre estes factores, destaca as dimensións das mallas empregadas na construción do copo, responsable da capacidade de selección de tallas do sistema de pesca. A correcta especificación dos tamaños de malla, polo tanto, é un dos apartados de maior importancia na xestión das pesqueiras. Idealmente, o mínimo tamaño de malla debería ser determinado en relación ao tamaño mínimo comerciabilizable das especies (He, 2006) dado que, unha elección pouco selectiva achega ao sector ante escenarios de insostibilidade dos stocks pesqueiros, e no caso contrario, os principais prexudicados a curto prazo son os pescadores, xa que tamaños de maia demasiado amplos puideran reducir as capturas de maneira insostible para as súas actividades económicas.

A pesqueira de arrastre de fondo obxecto de estudo neste caso práctico foi definida por Lart et al (2001), e, máis recentemente, por Castro (2007). O segundo autor, utilizando técnicas de clasificación multivariantes sobre as capturas totais das flotas, identifica a pesqueira a estudo como OTB-mixed (Bottom Otter trawl-mixed). As mareas de esta pesqueira teñen como obxectivo a captura combinada de distintas especies, entre as que destacan a Pescada (*Merluccius merluccius*), o Rapante (*Lepidorhombus boscii*), Rapes (*Lophius spp.*), Lírío (*Micromesistius potassou*) e Cigala (*Nephrops norvegicus*) (Castro, 2007). En xeral, os buques que compoñen esta pesqueira abarcan un rango de eslora comprendida entre os 21 e os 30 metros, e a duración das mareas están comprendidas entre 1 e 5 días (Pérez et al, 2007). As capturas obtidas en cada marea son mantidas en xeo e postas á venda sen eviscerar en portos do litoral norte español .

A orde [3] establece un plan de pesca para esta pesqueira, especificando no seu artigo terceiro, que a dimensión da malla será igual ou superior a 55mm. Estas dimensións son amplamente superadas nos buques que actualmente compoñen a flota, sendo as dimensións habituais comprendidas entre 60-69mm.

O caso práctico que a continuación se expón parte de unha acción piloto levada a cabo no 2007, que tivo como obxectivo avaliar medidas técnicas de conservación das especies nesta pesquería mixta, en concreto, a introdución de tamaños de mallas potencialmente máis selectivas que as que actualmente se atopan en uso.

## 3.2. Acción Piloto

A Acción Piloto de Pesca Experimental foi financiada pola Secretaría Xeral de Pesca Marítima e dirixida polo Instituto Español de Oceanografía (IEO) baseándose no proxecto promovido pola empresa armadora JOSÉ BALAYO, S.L. A experiencia tivo lugar a bordo do buque comercial JOSÉ BALAYO PORTELA (con porto base en Avilés).

Un dos obxectivos que perseguíu esta acción piloto, foi a avaliación, en termos de rendemento e selectividade das capturas, do funcionamento do arte de pesca estándar na pesqueira, ante un cambio no tamaño de malla dos copos:

- dende os 70mm, en adiante denotado por como  $C$ , malla utilizada habitualmente pola flota;
- ata os 80mm, en adiante denotado por  $T$ , malla test.

Para acadar os obxectivos marcados na Pesca Experimental propúxose o seguimento científico mediante a realización de 45 días de embarques de un observador científico, previamente seleccionado e formado para levar a cabo as mostraxes, e avaliar a correcta implementación do deseño experimental. O observador embarcouse nas mareas de pesca comercial comprendidas entre o 22 de Outubro e o 5 de Decembro do 2007. A duración do periodo de mostraxe foi de 33 días de pesca efectivos. A figura(3.1) mostra a zona de pesca da acción piloto, que se corresponde coas augas do litoral cantábrico, entre Cabo Peñas (Asturias) e Cabo Ajo (Cantábrica) ( División VIIIc da Comisión Internacional para a Exploración do Mar (ICES)).

## 3.3. Sistema de Pesca

O sistema de pesca utilizado na campaña denominase coloquialmente como “raspita” (figura(3.2)), un arte de arrastre que opera nos ambitos bentónicos e demersales (en contacto co fondo mariño), gracias a uns lastres denominados portas (figura(3.3)) ,que, conectados ao arte, manteñen ao conxunto en contacto co fondo durante o proceso de pesca. En faena, estes sistemas caracterízanse por unha ampla abertura horizontal, en detrimento das dimensións verticais, próximas aos 2 m. Tanto o deseño como a construción do arte utilizado son análogos aos utilizados de maneira comercial

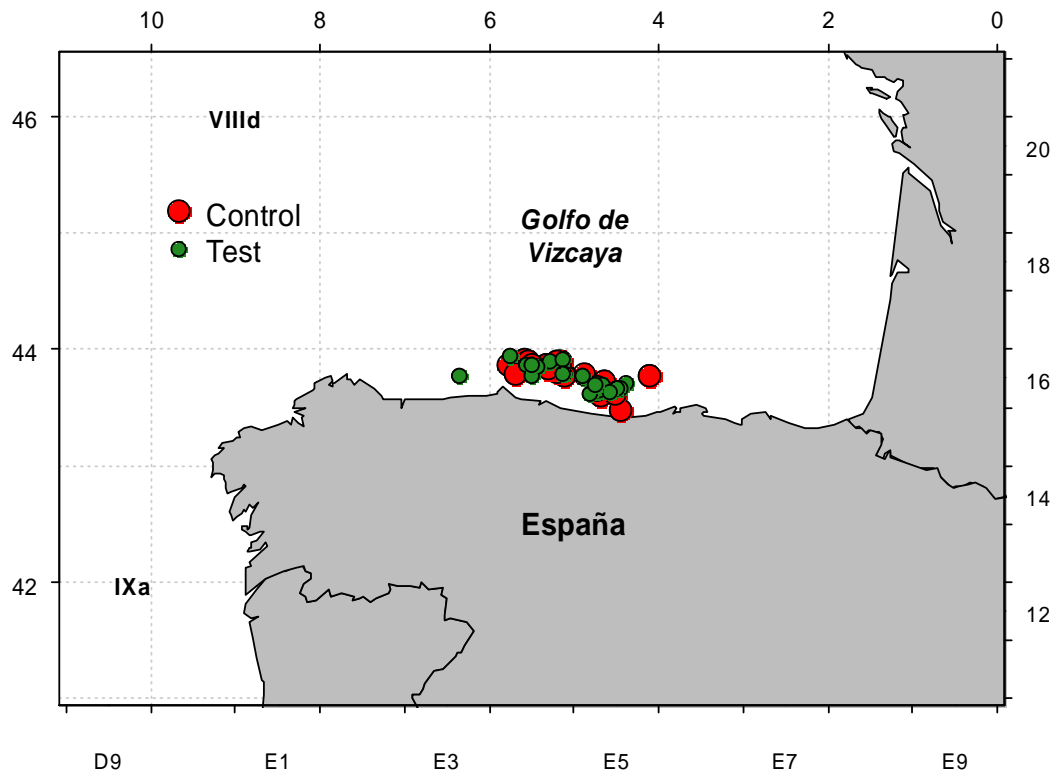


Figura 3.1: Zona de pesca e lances realizados utilizando os copos *C* e *T*.

na pesqueira. Este é un punto fundamental no deseño experimental, dado que, entre outros factores, as características diferenciais dos sistemas extractivos determinan o tipo e a cantidade de descartes xerados (Rochet e Trenkel, 2005). Polo tanto, a utilización de un arte diferente non permitiría extrapolar os resultados da campaña ao total da flota.

### 3.4. Deseño Experimental

Debido ás condicións de traballo e á mellor precisión na estimación dos parámetros de selectividade, optouse polo sistema de capturas con sobrecopo. Con este método, o copo queda incluído dentro de unha cuberta con malla de menor dimensión, deseñada para impedir que os individuos capturados polo arte, e filtrados a través das mallas do copo, queden en liberdade .En



definitiva, esta técnica permite coñecer a cantidade de captura filtrada a traves das mallas do copo.

Buscando unha boa calidade de datos, procurouse realizar as probas pesqueiras en zonas de alta abundancia das especies obxectivo dos estudos de selectividade (entre elas, Pescada, Lirio e Rapante).

Durante as mostraxes realizados en cada lance de pesca, o observador científico tivo a responsabilidade de recompilar información acerca da estrutura de tallas das especies máis importantes capturadas por cada especificación pesqueira (copos  $C$  e  $T$ ). Esta información foi obtida tanto sobre a captura retida no copo, como aquela que se filtrou ao sobrecopo. O observador puxo especial atención a que en ningún momento se producira a mezcla entre ambas porcións de captura, o cal puidera desvirtuar procesos analíticos posteriores.

### 3.5. Obxectivos

Os obxectivos específicos deste caso práctico son:

- Estimar as curvas de retención das especies Rapante, Pescada e Lirio según os copos  $C$  e  $T$ , utilizando os modelos descritos no capítulo 2.
- Utilizar as técnicas de inferencia presentadas no mesmo capítulo, co fin de avaliar posibles diferencias nas curvas de retención de  $C$  e  $T$

### 3.6. Análises

A metodoloxía estatística aplicada neste caso práctico, toma como base os procedementos expostos no capítulo 2, relativos aos modelos, estimación e inferencia de curvas. Os argumentos matemáticos específicos do caso práctico son desenrolados conxuntamente coa exposición dos resultados

#### 3.6.1. Estimación de Curvas de Retención

Sexa  $Z$  o factor que denota o copo utilizado nun determinado lance pesqueiro do presente caso práctico, con niveis  $C$  e  $T$  dados polo tipo de copo empregado, denotamos por  $Y$  ao número total de peixes que foron

retidos no copo  $Z_i$ , ao longo da acción piloto, e a  $U$  como o número de individuos filtrados a través das mallas dese mesmo copo. Denotaremos por  $p(l)$  a curva de retención do copo a estudio, que se corresponde coa probabilidade de que un peixe de talla  $l$  sexa retido no copo dado que entrou en el. En este marco, o número de peixes de talla  $l$  retidos no copo, condicionado ao total de capturas  $n(l) = Y(l) + U(l)$ , segue unha distribución binomial

$$Y(l) \sim \text{Binomial}(n(l), p(l)) \quad (3.1)$$

A función de probabilidade  $p(l)$ , que relaciona a proporción de retención en función das tallas, pode ser modelizado a través do modelo lineal loxístico tal que

$$p(l) = \frac{\exp(\beta_0 + \beta_1 l)}{1 + \exp(\beta_0 + \beta_1 l)} \quad (3.2)$$

Unha xeneralización do modelo anterior ben dado polo modelo GAM

$$p(l) = \frac{\exp(\beta_0 + f(l))}{1 + \exp(\beta_0 + f(l))} \quad (3.3)$$

sendo  $f(l)$  unha función suave e arbitraria.

En base a estes modelos de regresión levouse a cabo a estimación das curvas  $p(l)$  para para os copos  $C$  e  $T$ . As figuras 3.4,3.5,3.6,3.7,3.8,3.9 mostran ditas curvas xeradas polos métodos GLM e GAM, sobre as tres especies a estudio. Os suavizadores utilizados para as estimacións dos primeiros foron os P-splines (“ps”), e o grado de suavización foi estimado de maneira automática polo criterio UBRE (`mgcv`), se ben, ante a baixa suavización ocorrida para certos casos, decidiuse limitar os grados de liberdade a un máximo de 6. Como se esperaba, os resultados amosan que a talla é un factor significativo que condiciona a retención no copo dos individuos capturados. A flexibilidade propia dos GAM produciron unha visible mellora no axuste das curvas de retención aos datos. En concreto, estes modelos solucionaron o problema de axuste observado nos modelos GLM para os casos do Lirio en ambas copos, e a Pescada capturados polo arte T. O GLM estimou unha curva próxima á liña recta para o Rapante retido polo arte  $T$ , sen embargo, o modelo GAM conserva unha esperable curva monotónica crecente en forma de  $S$  para o mesmo caso. Os intervalos de confianza de cada unha das curvas foron xerados por Bootstrap ( $B = 1000$ ). En todos os casos os intervalos xerados foron moi

curtos, a excepción da curva de Rapante e Pescada no arte  $T$ . No primeiro caso, o Bootstrap ampliou os intervalos en clases de tallas altas, ante unha inesperada baixada na proporción de retención contida nos datos (talla 35). Na curva de retención da Pescada do arte  $T$ , o Bootstrap captou a alta variabilidade nas proporcións de retención das clases de tallas altas.

### 3.6.2. Contrastes de Interacción

O interese de este exercicio reside na posibilidade de inferir se un cambio de 10 mm. no tamaño dos copos  $C$  e  $T$  puido alterar o efecto que exerce a talla sobre a proporción de individuos retidos polo arte de pesca.

Sendo  $Z$  a variable categórica que inclúe aos dous tipos de copos,  $C$  e  $T$ , o posible efecto do copo na proporción de retidos en función da talla pode introducirse en (3.3) como:

$$p(Z, l) = \frac{\exp(\beta_0 + f(Z, l))}{1 + \exp(\beta_0 + f(Z, l))} \quad (3.4)$$

con

$$f(Z, l) = f^0(l) + \begin{cases} f^C(l) & \text{se } Z = C \\ f^T(l) & \text{se } Z = T \end{cases} \quad (3.5)$$

onde  $f^0$  denota o efecto principal da covariable talla, e  $f^C$  e  $f^T$  os efectos específicos da talla para os tipos de copo  $C$  e  $T$  respectivamente.

Asumindo o modelo anterior, contrastarase o modelo nulo

$$H_0 : f(l) = f^0(l) \quad \text{para todo } l$$

fronte ao modelo alternativo dado en (3.5). Con ese obxectivo empregouse o seguinte estatístico de contraste

$$T = \sum_{i=1}^n \left( \hat{f}^C(l_i) I_{\{Z_i=C\}} + \hat{f}^T(l_i) I_{\{Z_i=T\}} \right)^2$$

A distribución do estatístico  $T$  foi estimada mediante o procedemento Bootstrap exposta no Capítulo 2

As figuras 3.10, 3.11, 3.12, mostran graficamente o contraste de interacción, representando a diferenza vertical entre as curvas  $C$  e  $T$  e os intervalos

de confianza de tales diferencias. En cada unha das figuras se expón o p valor (distribución Bootstrap) obtido no test de hipótese. No caso do Rapante, atopamos un maior e inesperado poder de retención a favor do copo  $T$ , en clases de talla menores de 14 cm. Sen embargo, a capacidade de retención muda de tendenciá a partires de esa talla, en favor do copo con menor dimensión de malla ( $C$ ). O máximo diferencial a favor do copo  $C$  situase dentro das clases de tallas comerciais de esta especie ( $\sim 22$ cm), e a perda de capturabilidade do arte  $T$  respecto a  $C$  nestas clases de tallas e estimada en  $\sim 40\%$ . Ambas especificacións igualan o seu poder de retención na talla  $\sim 29$  cm, o que pode entenderse como o límite a partir do cal o tipo do copo deixa de ter un efecto significativo na retención de esta especie. O test de distribución Bootstrap rexeita a hipótese nula de igualdade das curvas de retención para estas especie.

As curvas de retención diferenciais das especies fusiformes (Lirio e Pescada) mostran unha tendencia semellante á anterior, dando maiores e inesperadas capacidades de retención ao copo  $T$  para as tallas menores. O cambio de tendencia prodúcese na talla  $\sim 24$  cm. No caso do Lirio, e na talla  $\sim 21$  cm. na Pescada. A retención de clases de talla comerciais de Pescada viuse moi afectada ante a utilización de  $T$ , con perdas de capacidade de retención máximas do  $\sim 40\%$  en tallas  $\sim 36$  cm., sen embargo, as perdas foron menores no caso do Lirio  $\sim 20\%$  (tallas  $\sim 27\%$ ). O test de distribución Bootstrap rexeita a hipótese nula de igualdade das curvas de retención para ambas especies.

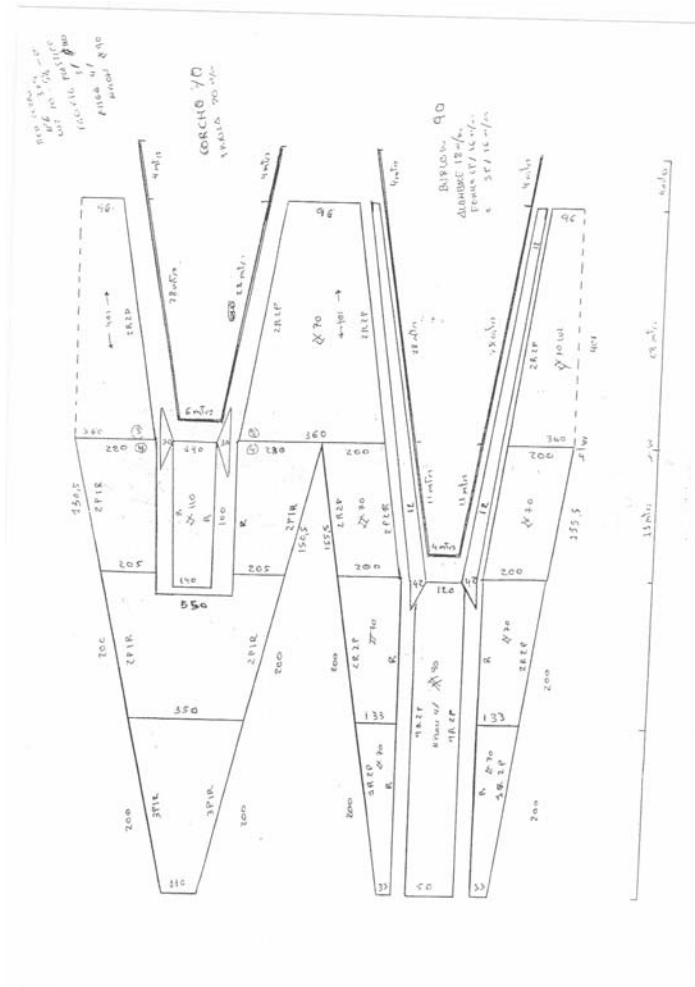


Figura 3.2: Plano clássico mostrando as especificações técnicas do arte de arrastre tipo “raspita”.



Figura 3.3: Porta convencional para a pesca de arrastre demersal.

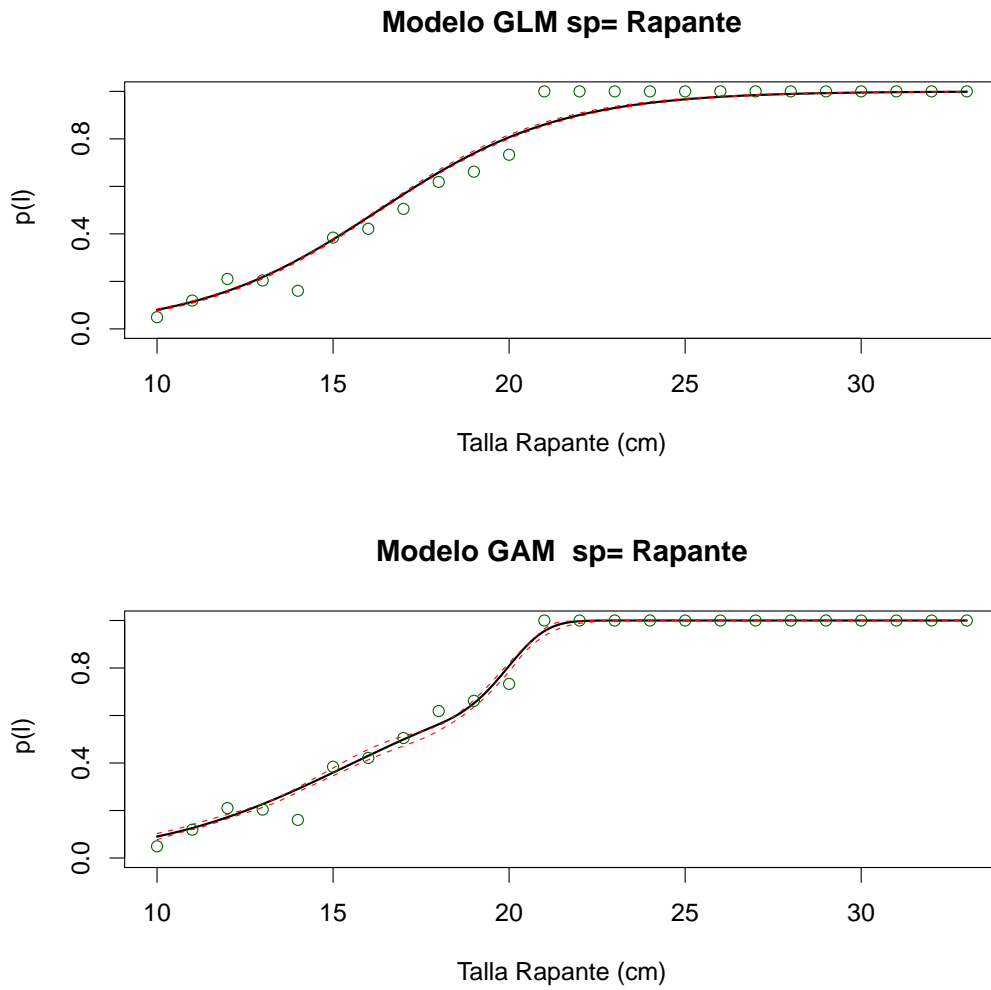


Figura 3.4: Curvas de retención para o Rapante, según modelos GAM e GLM para o arte *C*.

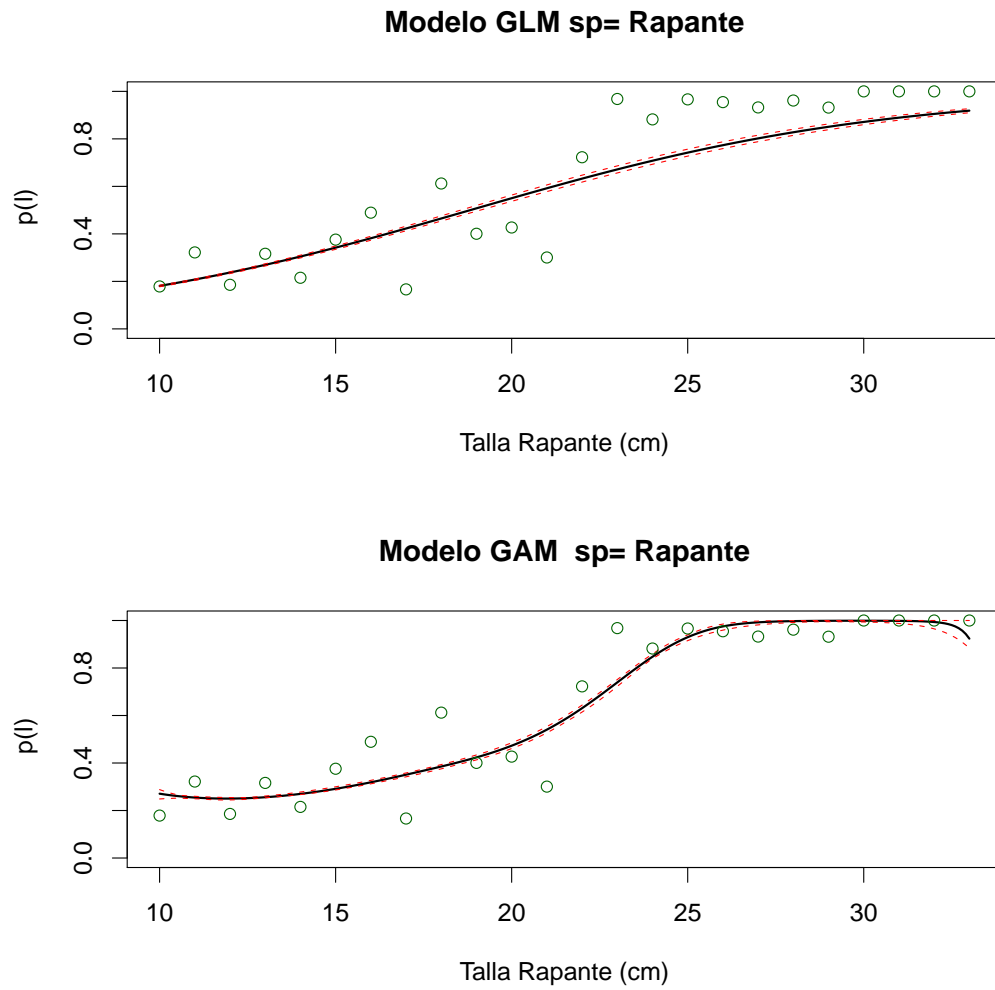


Figura 3.5: Curvas de retención para o Rapante, según modelos GLM e GAM para o arte  $T$ .



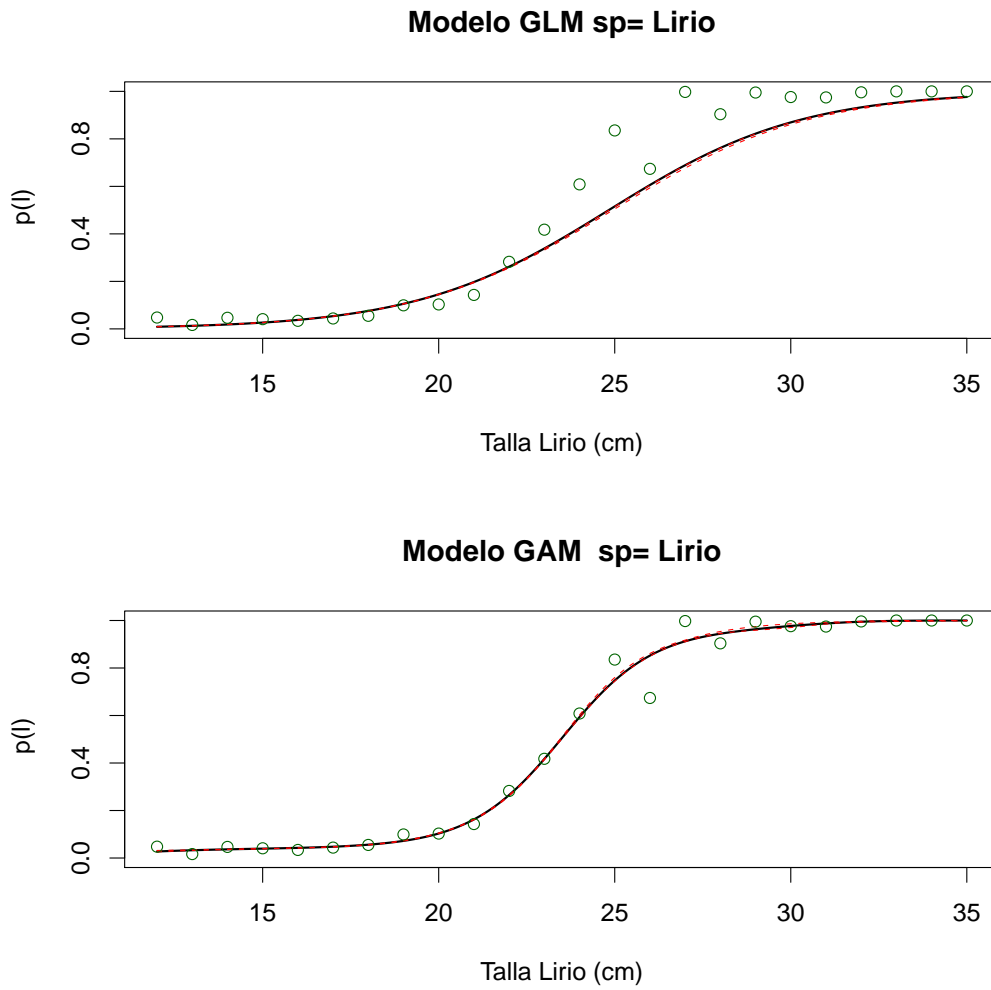


Figura 3.6: Curvas de retención para o Lirio, según modelos GLM e GAM para o arte  $C$ .

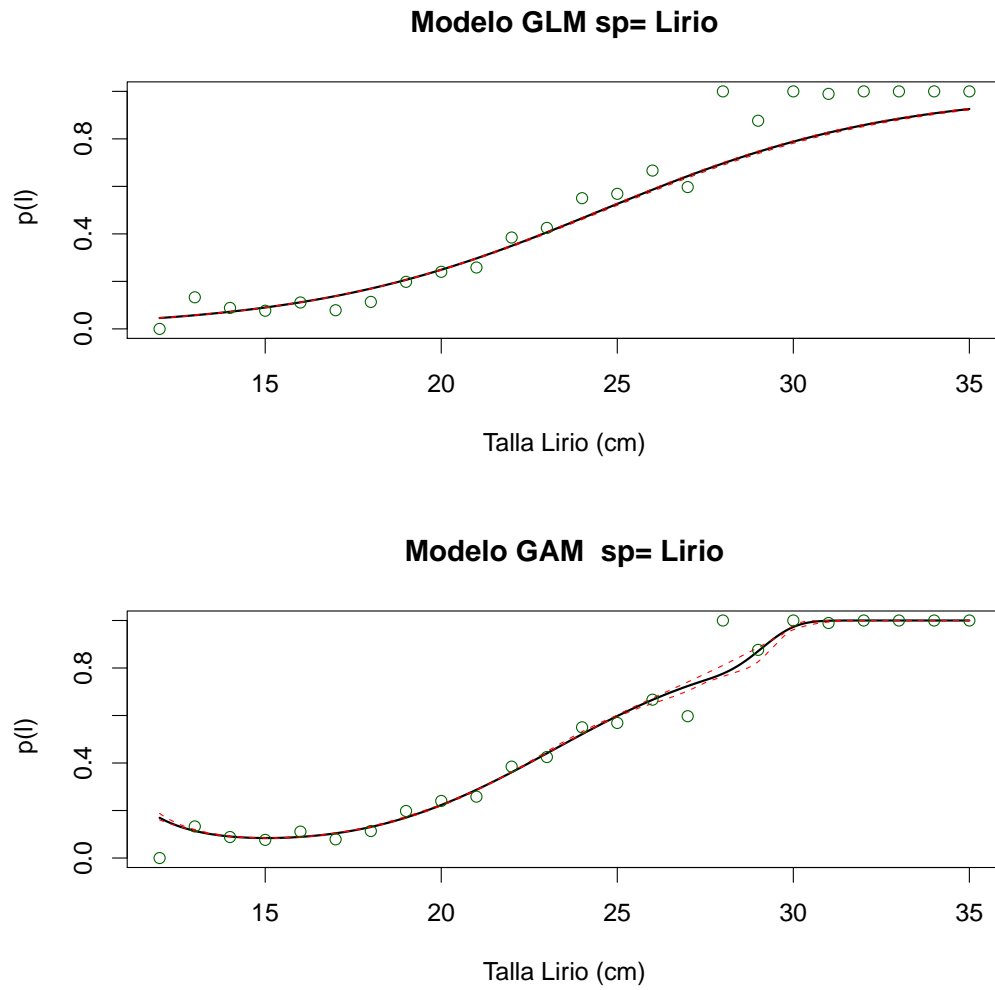


Figura 3.7: Curvas de retención para o Lirio, según modelos GLM e GAM para o arte  $T$ .

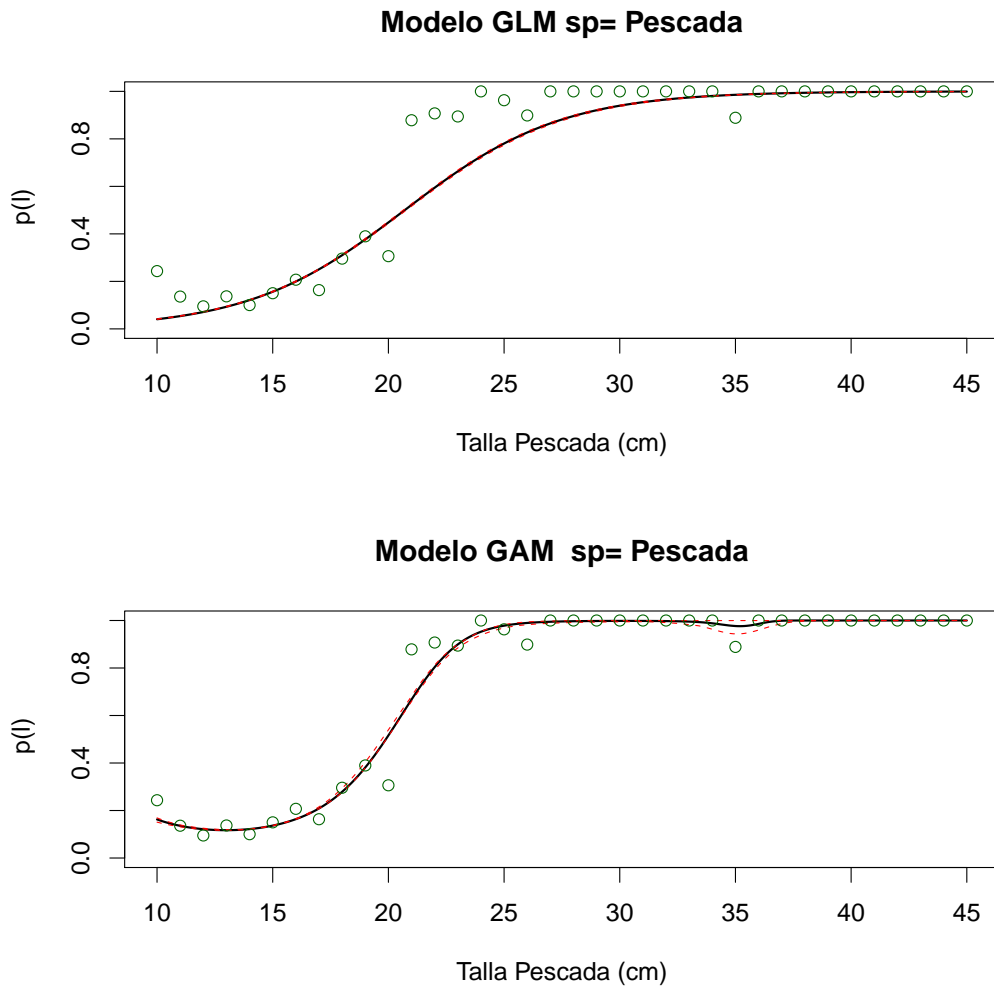


Figura 3.8: Curvas de retención para a Pescada, según modelos GLM e GAM para o arte  $C$ .

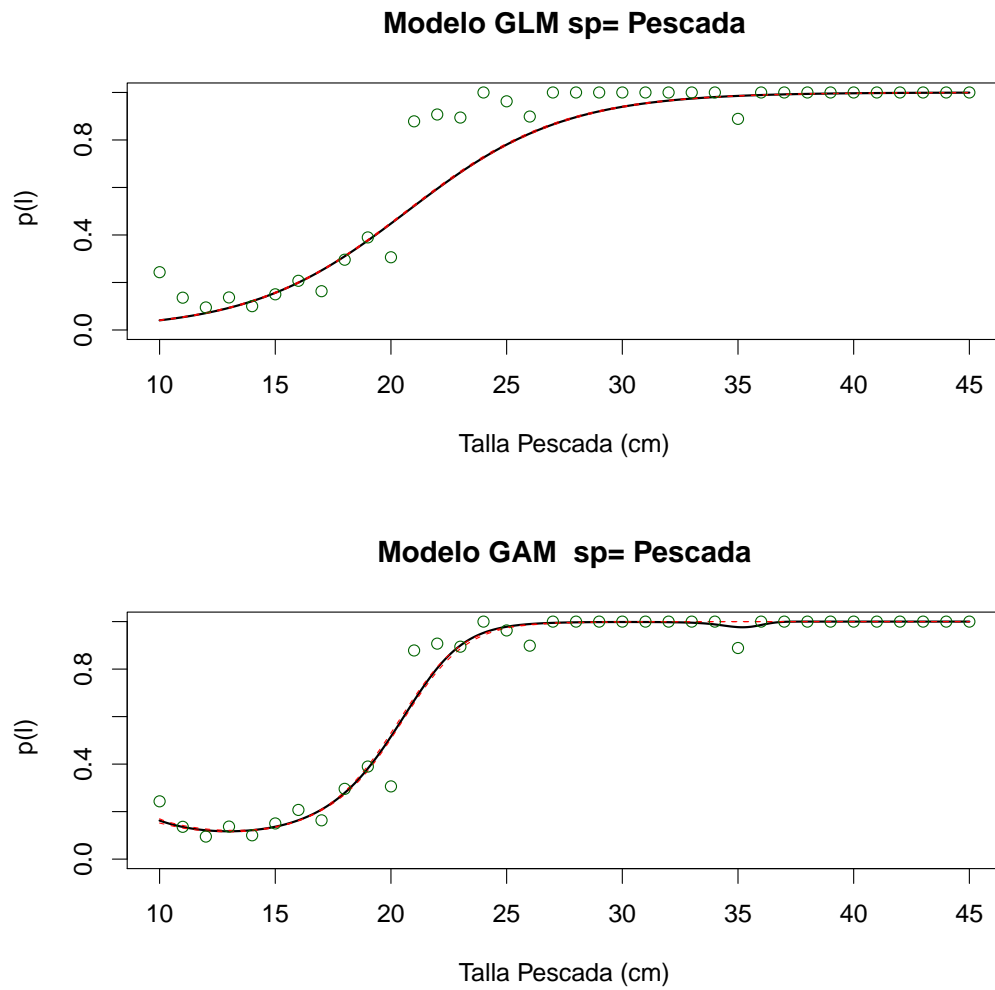


Figura 3.9: Porta convencional a Pescada, según modelos GLM e GAM para o arte  $T$ .

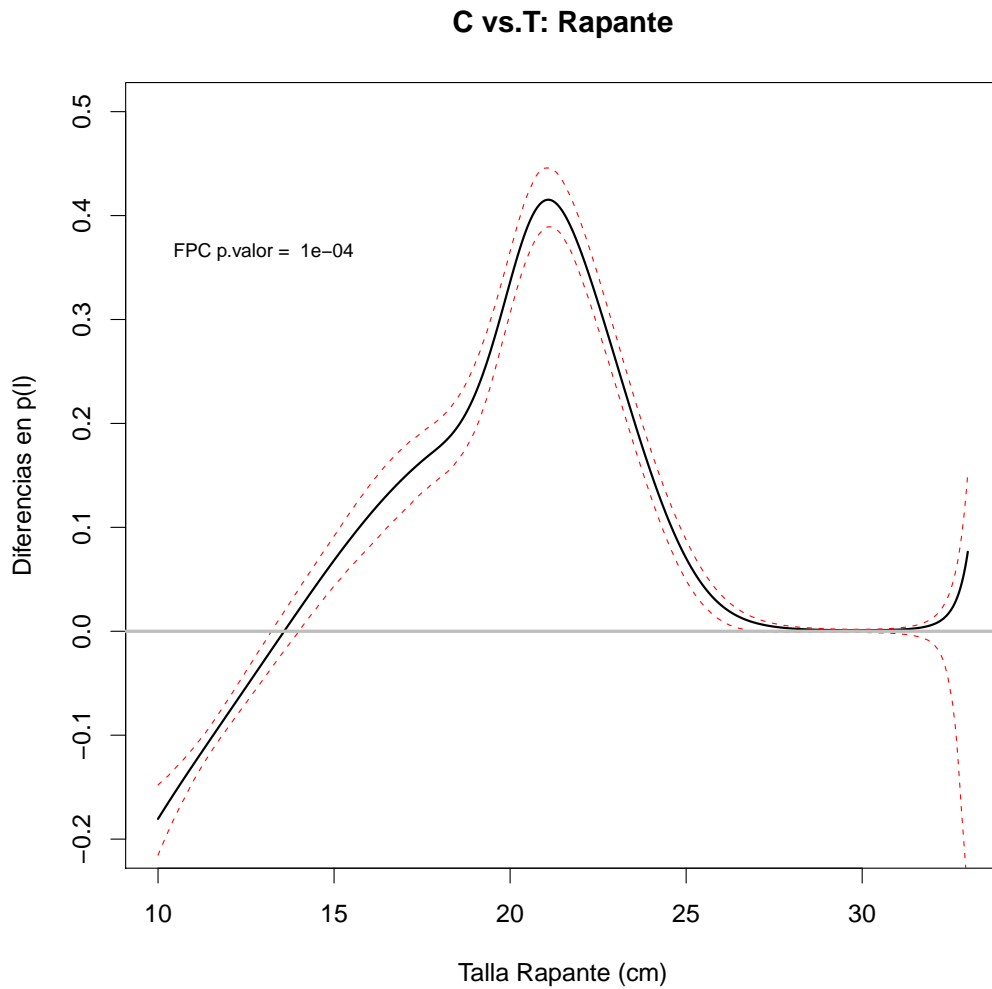


Figura 3.10: Diferencias e intervalos de confianza Bootstrap das diferencas entre as curvas de retención  $C$  e  $T$  estimadas para o Rapante. A presenza da curva diferencial na zona de valores positivos indica un maior poder de retención para o arte  $C$ . Os valores negativos nas clases de talla menores indican un maior e inesperado poder de retención de  $T$ . O p-valor ( $> 0,001$ ) indica diferencas significativas entre as curvas. As curvas de retención igualanse a partires da talla  $\sim 39$ .

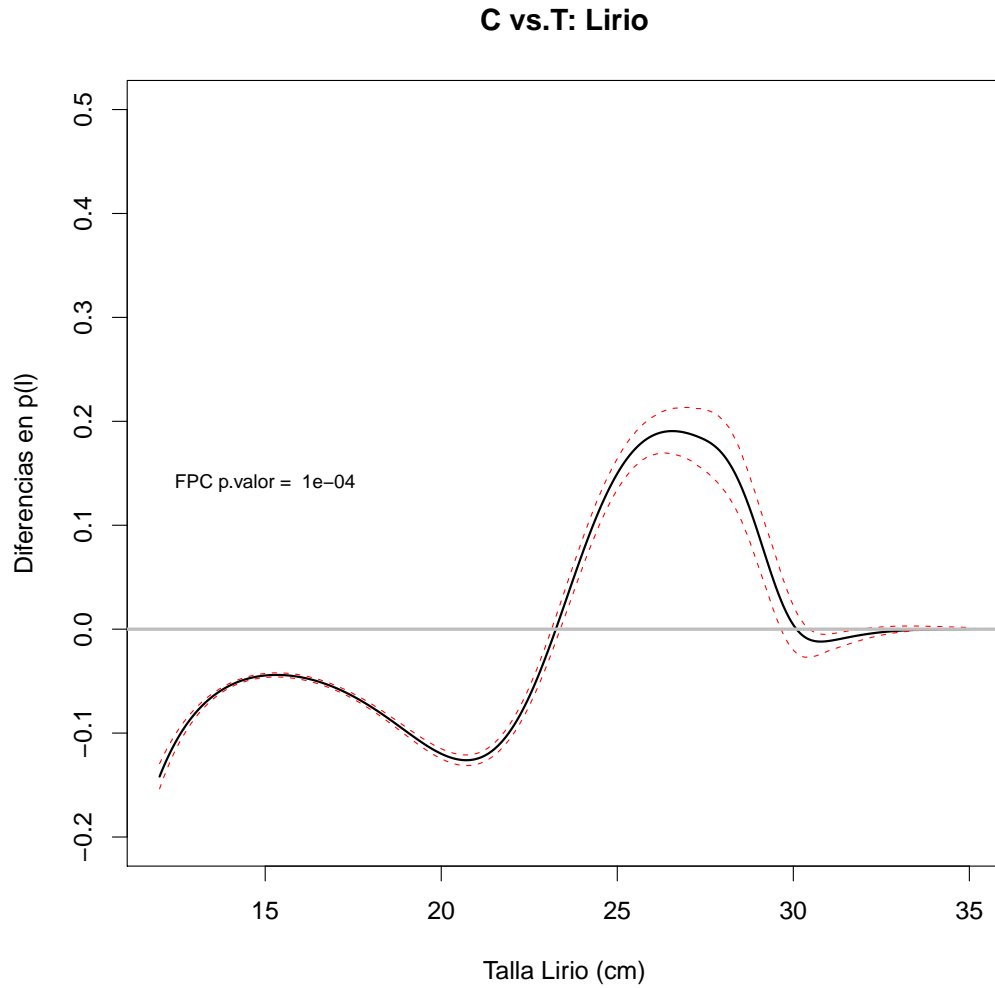


Figura 3.11: O Novamente  $T$  acaba maiores e inesperadas capacidades de retención en tallas menores. A presenza da curva diferencial na zona de valores positivos indica un maior poder de retención para o arte  $C$ . O p-valor indica diferenzas significativas entre as curvas ( $> 0,001$ ).

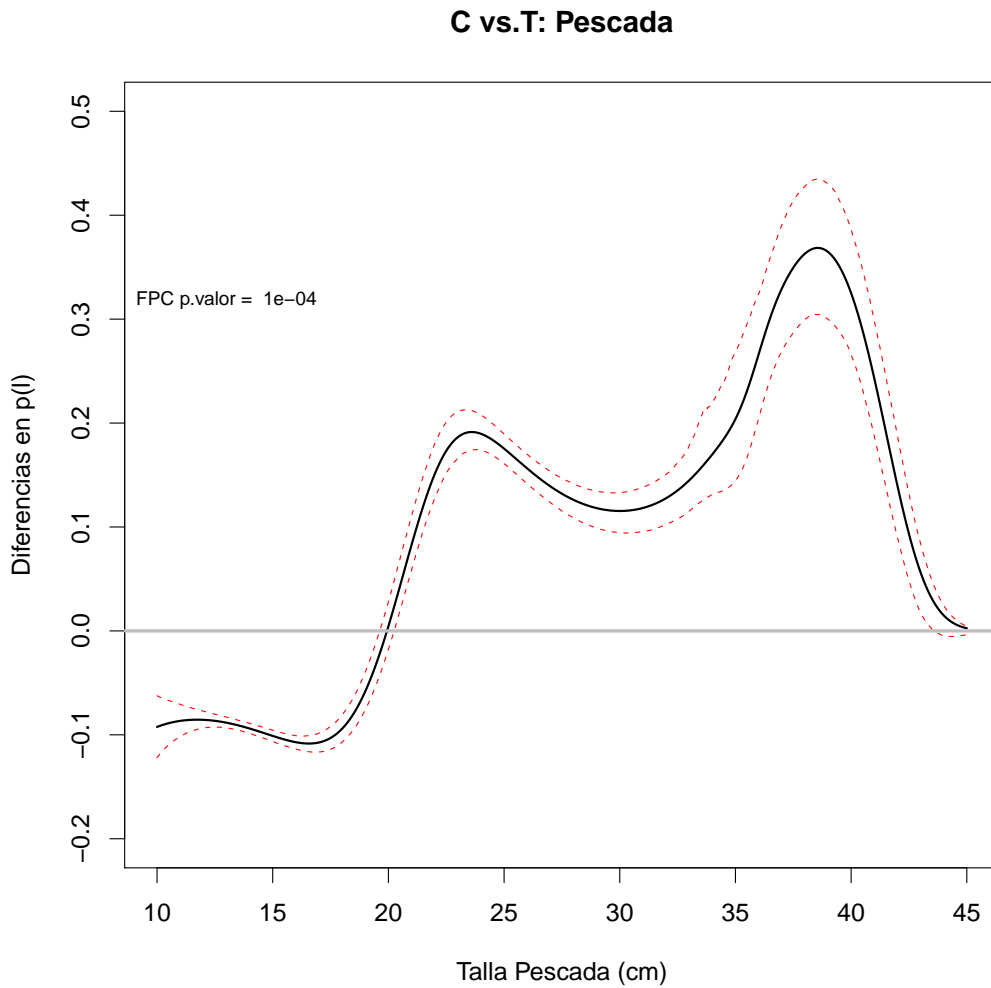


Figura 3.12: A curva diferencial e os intervalos de confianza para a Pescada indican que o copo  $C$  presentou maior poder de retención nun amplo rango de tallas. Como nos casos do Rapante e o Lirio, atopamos neste caso unha maior e inesperada capacidade de retención do arte  $T$  para as clases de talla menores da especie ( $p\text{-valor} > 0,001$ ).





# Capítulo 4

## Caso Práctico 2

## 4.1. Antecedentes

O Regulamento [4] da Comunidade Europea, establece medidas técnicas adicionais ao Regulamento [5], co fin de recuperar a poboación da Pescada (*Merluccius merluccius*) nas subzonas ICES (International Council for the Exploration of the Sea) III, IV, V, VI e VII e divisións ICES VIII a,b,d,e, mediante a imposición de normas que condicionan as actividades pesqueiras. Entre outros aspectos, esta lexislación trouxo consigo a adecuación da dimensión de malla, segundo as zonas de pesca e as flotas. Tal medida trata de diminuír a mortalidade de xuvenís, maximizando o escape de estas clases de idade unha vez capturados. En concreto, especificase a prohibición de operar con sistemas extractivos confeccionados con tamaños de malla inferiores a 100 mm, na área exposta na figura 4.1.

A baixa selectividade que caracteriza á maior parte das técnicas de pesca de arrastre demersal, xunto coa co-ocorrenza biolóxica das especies que son vulnerables a estes artes, imposibilita en moitos casos a caracterización monoespecífica dos seus obxectivos pesqueiros, dado que un considerable número de especies e clases de tallas poden ser retidas nun mesmo lance. Esta fenómeno habilita un escenario onde o pescador tratará de seleccionar as especies e individuos que maximicen o rendemento económico da súa actividade. Polo tanto, calqueira regulamentación de tipo monoespecífico, destinada a mitigar o impacto pesqueiro sobre unha determinada especie, pode ser de difícil aplicación neste tipo de pesqueiras mixtas (Murawski, 1983).

Este caso práctico xurde a raíz da aplicación da regulamentación arriba exposta, o que provocou un forte rexeitamento en pesqueiras caracterizadas por capturas de tipo mixto. Os prexudicados consideran a medida como pouco consecuenta coa natureza diversa das flotas, á vista de que algunhas flotas pesqueiras que operan na zona de exclusión, non valoran á pescada como o seu principal obxectivo pesqueiro-comercial, senón como unha especie de interese entre outras. Baixo esta problemática atopase a flota “Arrastre de rapantes Vi-Vii”, con base no porto de Vigo, e zonas de pesca incluídas nas áreas ICES VI y VII. Esta flota ten como especies obxectivo, ás dúas especies de Rapante (*Lepidorhombus whiffiagonis* e *Lepidorhombus boscii*) presentes na área, que, xunto ao Rape (*Lophius spp.*), representan entre o 65-80% do valor anual desembarcado (Lart et al, 2002), a Pescada, neste caso, queda enmarcada nun segundo plano de interese económico. O sector expón que a nova regulamentación (incremento de 80 a 100mm) merma a capturabilidade de outras especies diferentes á pescada, en concreto, os

pescadores detectan unha forte diminución de capturas de Rapante, feito que pon en perigo a viabilidade económica das súas actividades pesqueiras.

## 4.2. Acción Piloto

Tratando de analizar esta problemática, a Secretaría Xeral de Pesca Marítima financiou no 2006 unha campaña experimental denominada “*Pesca Experimental RAI-AP-10/2006 Dirixida á Adecuación do Tamaño de Malla na Pesqueira de rapante Mediante Arrastre de Fondo en NEAFC*”, baixo a dirección do Instituto español de Oceanografía (C.O. Vigo). A experiencia trataría de contrastar analiticamente os argumentos dos pescadores en canto ás hipotéticas perdas provocadas polo cambio de regulamentación. Polo tanto, o primeiro obxectivo foi a comparación das capturas obtidas co copo que viña utilizando a flota (tamaño de malla de 80 mm.), respecto a aquelas obtidas co tamaño de malla selectivo (100 mm.). As probas pesqueiras foron levadas a cabo polo *B/P Pescaberbes III*, barco que pertence á Unidade Operacional en cuestión, con base no porto de Vigo, que faena habitualmente na Sub-área de pesca ICES suxeita ás restricións arriba expostas. Previo inicio de campaña, foron seleccionados dous observadores pesqueiros, responsables das mostraxes a bordo. No proceso de selección valorouse fundamentalmente a experiencia dos mesmos en tarefas de mostraxe. Unha vez seleccionados, levouse a cabo unhas xornadas de formación onde se puxo en común tanto o deseño experimental, como as mostraxes planificadas para a campaña.

A experiencia tivo unha duración de 36 días, distribuídos ao longo de dúas mareas continuas, dende o 6 ata o 24 de Outubro a primeira, e entre o 2 e o 20 de novembro do 2006 a segunda, e executados en zonas próximas á área de restrición 4.1. O resultado foi a consecución de 52 lances efectivos con cada un dos copos. Unha vez a bordo, os observadores levaron a cabo o control tanto dos resultados pesqueiros obtidos en cada lance, como das variables suxeitas a cada proceso pesqueiro.

## 4.3. Sistema de Pesca

O sistema de pesca utilizado foi o mesmo que o empregado no caso práctico 1. Dous destes artes de pesca estiveron operativos ao longo da campaña de maneira simultánea, variando exclusivamente no copo fixado á súa parte

posterior.

#### 4.4. Deseño Experimental

O método de mostraxe seleccionado nesta experiencia denomínase como “lances alternos” (Anónimo, 1996). Este método foi seleccionado entre outras posibilidades, dadas ás difíciles condicións meteorolóxicas que habitualmente imperan na zona de pesca, o que puidera xerar dificultades á hora de executar a metodoloxía do **caso práctico 1**. Neste deseño experimental, as pescas lévanse a cabo de maneira sucesiva, utilizando de maneira alternativa o copo control (C = 80mm.) e o copo test (T=100 mm.).

A idea principal desta metodoloxía reside en que, en teoría, cando os distintos copos son utilizados nunha mesma zona e en tempos sucesivos, accederán ao mesmo recurso pesqueiro, e polo tanto, as súas capturas serán comparables (Fryer et al., 2003). A maior desvantaxe deste método é a alta variabilidade non controlada que pode existir dentro de unha mesma serie de dous lances, o que fai imprescindible elevar o número de probas para acadar certa validez estatística (Madsen 2007). Tendo en conta estes condicionantes, acordouse que a toma de decisións en torno á execución dos lances correra a cargo dos observadores, únicos tripulantes capacitados para valorar, tanto o correcto axuste do plan de mostraxes e deseño experimental, como a mínima desviación respecto as condicións de pesca comercial. Previamente ao inicio da campaña, acordouse manter estabilizado o maior número de variables controlables, especialmente dentro de cada serie, polo tanto, unha serie identificaríase como válida cando os lances que engloba foran executadas a unha mesma profundidade, luminosidade, área xeográfica, tempo sucesivo, etc...

Co fin de emular as pescas comerciais, a dirección do IEO decidiu ceder a captura obtida en cada pesca para a súa comercialización dentro da rutina habitual do buque.

#### 4.5. Obxectivos

- Adaptar as técnicas especificadas no **capítulo 2**, e postas en práctica no **caso práctico 1**, a un novo deseño experimental de acción piloto.

- Contrastar a hipótese plantexada polo sector, en canto á perda de poder pesqueiro das especies obxectivo da UO Rapanteira de Vigo, dado o cambio de tamaño de malla disposto pola lexislación vixente.

## 4.6. Análises

A argumentación matemática deste caso práctico trata de asumir os obxectivos plantexados e o novo contexto experimental. De este xeito, en lugar de extraer unha curva de retención por cada unha das especificacións **caso práctico 1**, xeramos unha única curva, a cal tratará de modelar o ratio relativo de capturas entre ambos artes de pesca. Esta metodoloxía, denominada “Comparación de capturas”, ´e utilizada tanto para comparar artes de pesca comerciais (Holst e Reville, 2009), como para a calibración de buques oceanográficos (Fryer et al., 2003).

### 4.6.1. Comparación de Capturas

Denotamos a  $Y$  como o número de individuos de unha especie que son capturados ao longo da acción piloto co copo  $T$ , e denotamos a  $U$  como o número observado en  $C$ .

Neste caso práctico, adaptamos a argumentación do primeiro, expresando que o número de individuos de talla  $l$  atopados no copo  $T$ , condicionando ao total de capturas para esa talla  $n(l) = Y(l) + U(l)$ , son binomialmente distribuídos dacordo a

$$Y(l) \sim \text{Binomial}((n(l), p(l)))$$

Polo tanto, o ratio relativo de capturas de  $T$  respecto a  $C$  será neste caso a resposta que se pretende modelar, utilizando os mesmos GLM e GAM do **caso práctico 1**.

### 4.6.2. Contraste de Ratios de Captura

A hipótese nula a contrastar neste caso práctico foi:

$$H_0 : p(l) = 0,5$$

fronte a alternativa

$$H_1 : p(l) \neq 0,5$$

As figuras 4.4, 4.5 e 4.6 mostran os resultados do exercicio de comparación de capturas. Na marxe superior esquerda de cada unha de elas presentase o p-valor do contraste proposto. A utilización do arte con tamaño de malla Test(100mm.) implica unha importante diminución nas capturas de Rape, en comparación ás obtidas pola malla Control (80mm.). A curva GAM, xunto cos intervalos de confianza xerados por Bootstrap (95 %) estimaron que tales perdas son mínimas ( $\sim 5\%$ ) en clases de talla altas  $\sim 40\text{cm.}$ , mentres que a máxima diferenza ( $\sim 20\%$ ) queda localizada en tallas de  $\sim 25\text{cm.}$ , que se corresponden con individuos de alto interese económico para a flota. A perda de poder pesqueiro non foi tan claro no caso da Pescada, se ben o contraste rexeitou a hipótese de igualdade. O ratio de capturas obtido na acción piloto presenta unha alta variabilidade ao longo do rango de tallas, e a curva estimada non presenta unha tendencia clara. Según a estimación, o copo *C* obtivo mellores resultados no intervalo comprendido entre as tallas 21 e 40cm. Os ratios de captura relativa igualáronse en tallas próximas aos 60cm.. As mínimas diferenzas entre as dúas especificacións pesqueiras tiveron lugar comparando as capturas efectuadas sobre o Rape. A curva estimou para esta especie unha mínima pero significativa perda de capturabilidade do arte *T* en un rango de tallas comprendido entre os  $\sim 21$  e os  $\sim 50\text{cm.}$



Figura 4.1: Zona de exclusión de actividades pesqueiras con tamaños de malla menores de 100mm de luz.



Figura 4.2: Captura retida nun lance pesqueiro con arte de pesca pouco selectivo, habitualmente utilizados nas pesqueiras mixtas.



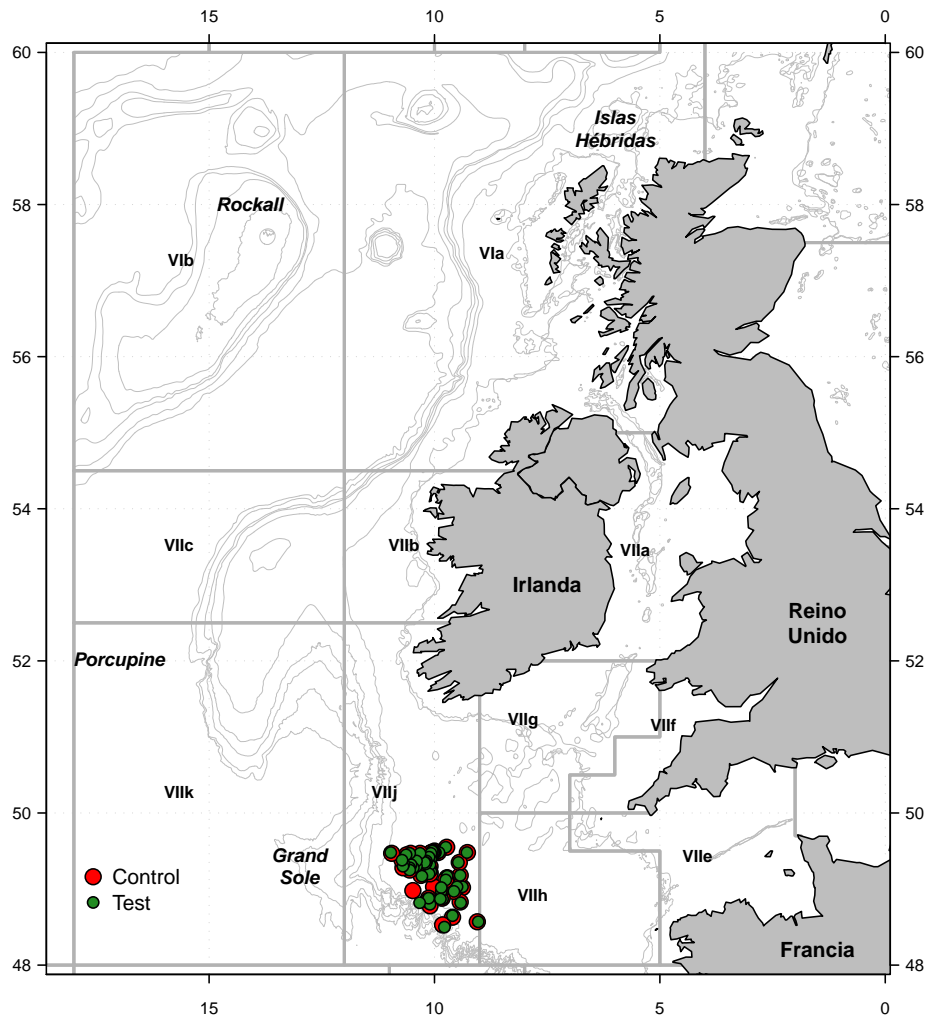


Figura 4.3: Zona de pesca e lances realizados utilizando os copos *C* e *T*

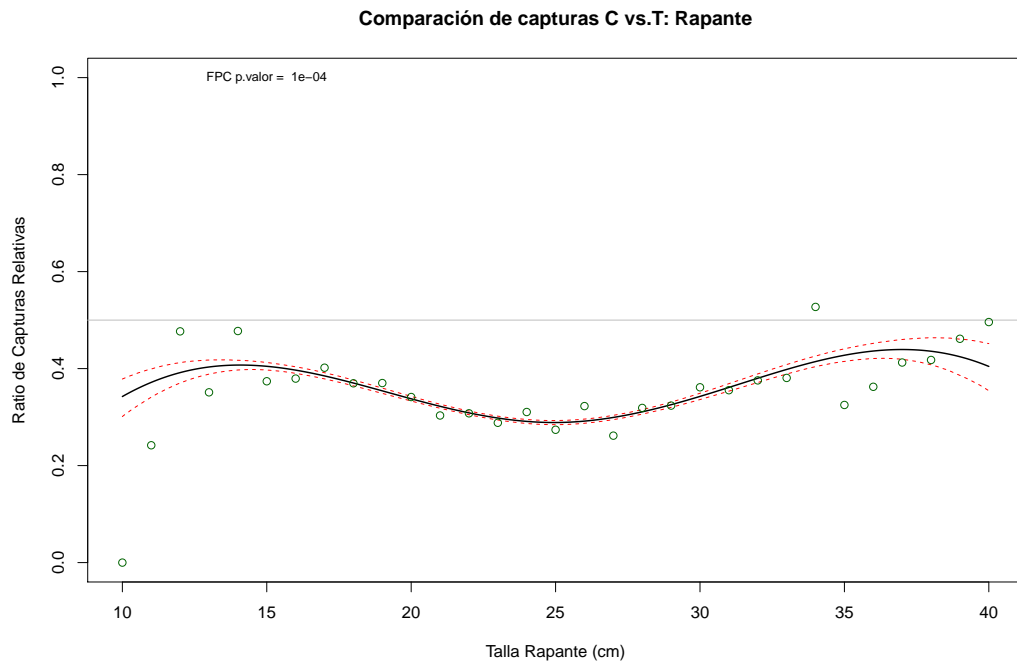


Figura 4.4: Curva media de ratio de capturas relativas e intervalos de confianza Bootstrap (95 %) para o Rapante. A liña gris horizontal situase no ratio 0.5, marcando situacións de igualdade de capturas. A perda de capturabilidade por parte de  $T$  é continua ao longo do rango de tallas.

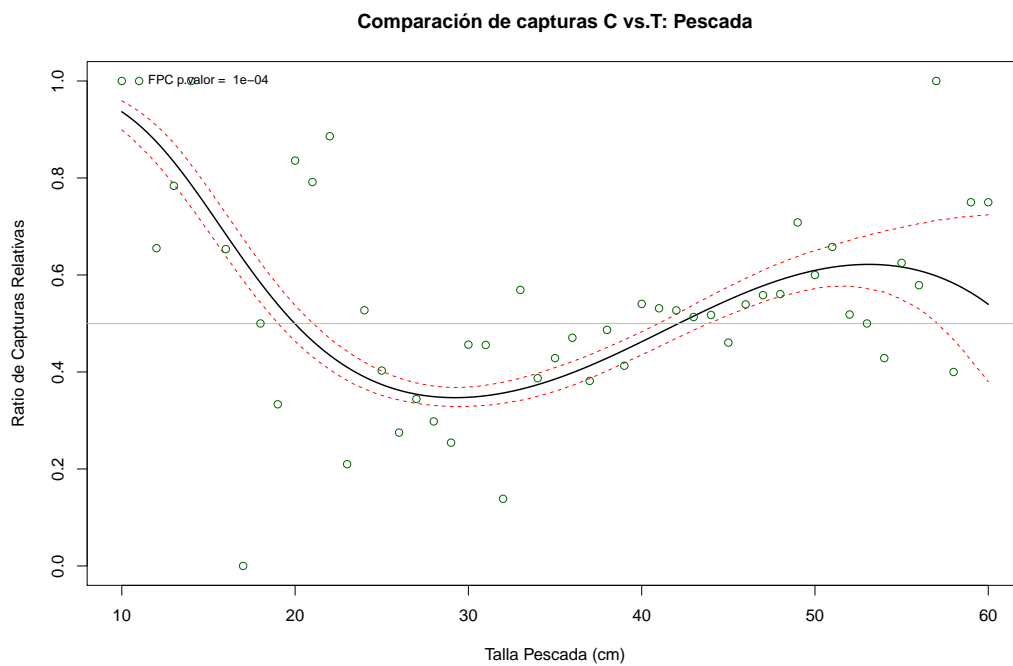


Figura 4.5: Curva media de ratio de capturas relativas e intervalos de confianza Bootstrap (95 %) para a Pescada. A curva non segue un patrón claro, oscilando en torno á liña de referencia.

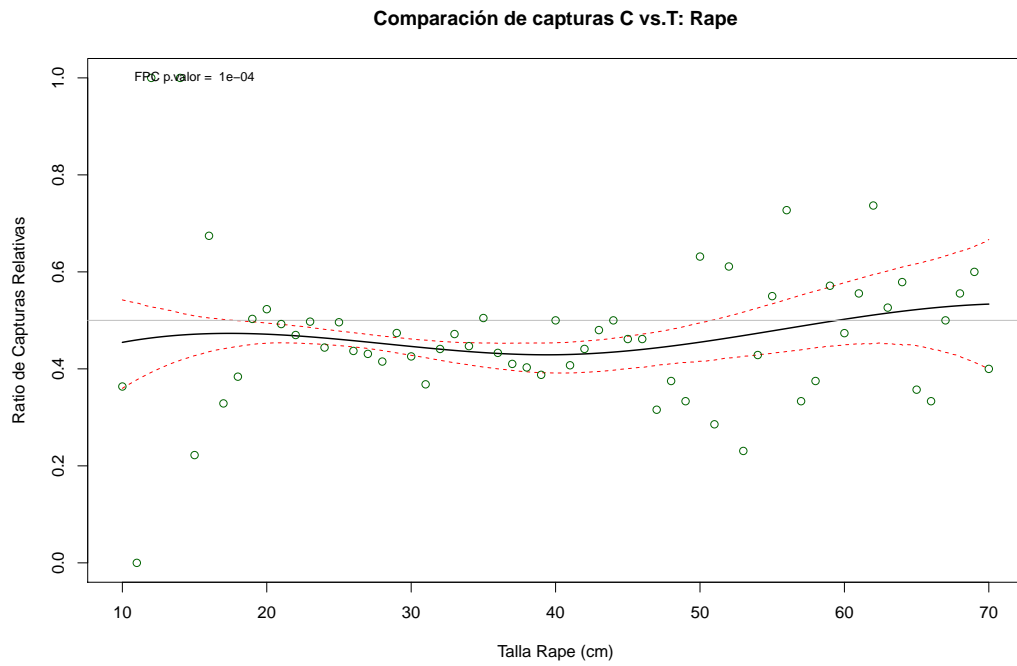


Figura 4.6: Curva media de ratio de capturas relativas e intervalos de confianza Bootstrap (95 %) para o Rape. A estimación captura a baixada observada no ratio relativo, en clases de tallas intermedias.

# Capítulo 5

## Discusión

O primeiro enfrontamento aos casos prácticos aquí expostos (Pérez et al., 2005; Pérez et al., 2007), deixou nos responsables das análises, a sensación de que os modelos paramétricos utilizados (GLM) eran demasiado ríxidos como para afrontar de maneira exitosa a modelación dos datos xerados nas accións piloto. Esta mesma percepción é compartida por distintos autores que, de maneira progresiva, foron aportando outras alternativas aos modelos paramétricos. Por exemplo, Millar (1993) modelou os procesos selectivos en dragas empregadas para o marisqueo de vieiras, atopando que as estimacións paramétricas non se axustaban adecuadamente aos seus datos. A alternativa empregada por este autor foi a estimación de curvas de regresión isotónicas. Munro e Somerton (2001) incorporaron os cubic smoothing splines para a estimación de curvas de selectividade. Os autores defenden a elección de este método argumentando a facil comprensión do proceso de selección do parámetro de suavización; en concreto, a Validación Cruzada (CV) foi utilizada neste caso para a determinación automática deste parámetro.

O presente traballo emprega con éxito os modelos de regresión spline penalizada como alternativa analítica para a estimación de curvas paramétricas de retención (**caso práctico 1**), e comparación de capturas (**caso práctico 2**). Sen embargo, o grado de suavización estimado de maneira automática polo método UBRE foi menor da que nun principio se esperaba. Fryer et al. (2003) argumentan que o grado de suavización é unha tarefa de sentido común e coñecemento previo do material a analizar. Na práctica, dito autor levou a cabo a elección do grado de suavización según procesos de proba-erro, valorando o correcto axuste dos datos, mediante a observación dos gráficos

de residuos. O autor empregou entre 2 a 4 grados de liberdade para a estimación das curvas de selectividade. No presente traballo mantívose constante o tope máximo de 6 grados de liberdade. Os resultados en canto á suavización son satisfactorios, se ben sería interesante investigar a raíz do problema de sobreaxuste observado para algúns casos aquí expostos.

O bootstrap é un procedemento habitualmente utilizado para incorporar a variabilidade natural das curvas de retención. Millar (1993), e Munro e Somerton (2001), utilizaron o mesmo enfoque que aquí se presenta, isto é, agrupando os datos de distintos lances en unha única matriz, axustando unha curva non paramétrica, e valorando a súa precisión mediante Bootstrap. Sen embargo, esta metodoloxía non é a única empregada na actualidade. A introducción dos modelos mixtos (Millar, 1993) foi rapidamente asimilada como unha alternativa metodolóxica, posta en práctica en diversos traballos.

Ambas alternativas son vistas como complementarias (Millar e Fryer., 1999), e a elección de un método ou outro segue un roteiro de preferencias persoais, máis que a outros aspectos de tipo funcional. O punto a favor máis destacado para o método Bootstrap, é que non require de asuncións teóricas previas, ás cales si deben axustarse os modelos mixtos. Sen embargo, estes últimos ofrecen a posibilidade de formular test de hipótesis sobre un contexto natural (non existe ningunha agregación de datos e polo tanto utiliza a variabilidade real observada entre lances). O bootstrap ao contrario, traballa con datos agregados, e a curva media é estimada en base á distribución conxunta de individuos retidos ao longo de  $n$  lances; este feito dificulta polo tanto o contraste de hipóteses que poideran xurdir neste tipo de traballos (Fryer et al., 2003). Os intervalos de confianza arroxados polo método bootstrap no presente traballo foron máis estreitos do que se poideran esperar a priori. Unha das posibles razóns sexa a agregación dos datos de entrada, coa conseguinte perda de variabilidade natural entre os lances. Por outro banda, a rixidez dos p-splines puidera levar a que as curvas xeradas en base as Pseudoreplicacións non variaran excesivamente. Esperase que futuros estudos de simulación, a partires dos datos analizados poidan dar luz a estas cuestións abertas.

Os resultados das estimacións de curvas de selectividade no caso **práctico 1**, son un claro exemplo da incertidume asociada aos procesos selectivos. A selección de tallas por parte do copo pode estar condicionado a moitos

factores (ademais do tamaño de malla) actuando no momento das pescas. Entre outros, o tipo e volume de captura que acompaña ás especies que se pretenden estudar, pode condicionar en gran medida o grado de abertura das mallas (Herrmann et al., 2006). O'Neill (2006) estuda as posibles interaccións entre o estado climático do mar, e a capacidade selectiva dos artes de pesca, atopando certa relación positiva entre ambos, e decir, canto peores son as condicións climáticas que atopa o arte de pesca no fondo do mar, maior batemento do copo, e polo tanto, maiores posibilidades de filtrado. A investigación destes e outros factores aportarán nun futuro interesantes resultados en canto aos procesos selectivos que teñen lugar nas flotas do litoral Atlántico da península ibérica.

O **caso práctico 2** aporta importantes resultados que respaldan os argumentos dos pescadores. A utilización de un tamaño de malla de 100mm. exerceu un forte efecto negativo sobre o rendemento das pescas. Sen embargo, existen algúns indicios que evidencian que outras causas, distintas ao propio proceso de filtración (selección pasiva), puideron incidir neste cambio de comportamento pesqueiro. En primeiro lugar, non é creíble que a diminución da capturabilidade do Rape sexa obra da selección pasiva das mallas. Debido ás especiais características morfolóxicas de esta especie, non se contempla que cambios de tamaños de malla poidan diminuír a retención de individuos da especie (Maartens, 2001). En segundo lugar, enquisas posteriores realizadas sobre patróns de dita pesqueira, coinciden en sinalar que o aumento do tamaño de malla produciu problemas de fixación do arte (ver figura 1.2) ao fondo mariño, e que estes problemas repercutiron negativamente na accesibilidade aos recursos pesqueiros. Mesmos argumentos foron testados en casos semellantes (Anónimo, 1996; Millar e Walsh, 1992), e os resultados destas investigacións parecen confirmar as sospeitas dos pescadores. Conversacións coa industria de manufacturación de redes, aclaran que o aumento do tamaño de malla pode producir unha serie de alteracións hidrodinámicas no interior do arte (vexase Herrmann et al., 2006), desembocando, en última instancia, na diminución do poder de barrido sobre o sustrato. Este fenómeno aumenta as posibilidades de que especies que viven sobre o sustrato mariño (como é o caso do Rapante e o Rape), sexan sobrepassados polo arte, sen ser barridos ao seu interior (Ryer, 2007). Esta situación, sen embargo, non terá efecto algún na capturabilidade de especies que, como a merluza, viven na columna de auga. Este enfoque pode dar unha explicación máis coherente á diminución xeralizada observada neste caso práctico para os Rapes e os Rapantes, incluso

en clases de tallas onde non se esperaba perda algunha por filtraxe a traves das mallas.



# Capítulo 6

## Bibliografía

- Alverson, D.L., Freeberg, M.H., Murawaski, S.A., Pope, J.G., 1994. A Global Assessment of Fisheries Bycatch and Discards. FAO Fisheries Technical Paper N° 339. FAO, Rome, p. 235.
- Anónimo, 1996. Manual of methods of measuring the selectivity of towed fishing gears. ICES Cooperative Research Report N° 215. ICES, Copenhagen, p. 125.
- Anónimo, 2004. The Nephrops fisheries of the Northeast Atlantic and Mediterranean - A review and assessment of fishing gear design. ICES Cooperative Research Report N° 270. ICES, p. 214.
- Beutel, D., Skrobe, L., Castro, K., Ruhle Sr, P., Ruhle Jr, P., O'Grady, J., Knight, J., 2008. Bycatch reduction in the Northeast USA directed haddock bottom trawl fishery. Fisheries Research 94, 190-198.
- Castro, J., 2007. Identification and segmentation of mixed-species fisheries operating in the atlantic iberian peninsula waters. p. 220.
- Catchpole, T., Revill, A., Dunlin, G., 2006. An assessment of the Swedish grid and square-mesh codend in the English (Farn Deeps) Nephrops

- fishery. Fisheries Research 81, 118-125.
- Catchpole, T.L., Frid, C.L.J., Gray, T.S., 2005. Discards in North Sea fisheries: causes, consequences and solutions. Marine Policy 29, 421-430.
- Cox, D.R., 1970. *The analysis of binary data*. Methuen, London
- Efron, B., 1979. Bootstrap methods: another look at the jackknife. The Annals of Statistics 7, 1-26.
- Efron, E., Tibshirani, R.J., 1993. *An introduction to the Bootstrap*. Chapman and Hall, London.
- Enever, R., Reville, A., Grant, A., 2009. Discarding in the North Sea and on the historical efficacy of gear-based technical measures in reducing discards. Fisheries Research 95, 40-46.
- FAO, 1996. Technical Consultation on Reduction of Wastage in Fisheries. FAO Fisheries Technical Paper N° 547. FAO, Rome.
- Ferno, A., Olsen, S., 1994. Marine fish behaviour in capture and abundance estimation, Oxford,.
- Fryer, R., Zuur, A., Graham, N., 2003. Using mixed models to combine smooth size-selection and catch-comparison curves over hauls. Canadian Journal of Fisheries and Aquatic Sciences 60, 448-459.
- Fryer, R.J., 1991. A model of between haul variation in selectivity. ICES Journal of Marine Science 48, 281-290.
- Gabr, M., Fujimori, Y., Shimizu, S., Miura, T., 2007. Behaviour analysis of undersized fish escaping through square meshes and separating grids in simulated trawling experiment. Fisheries Research 85, 112-121.
- Glass, C., 2000. Conservation of fish stocks through bycatch reduction: a review. Northeastern Naturalist 7, 15.
- Gray, C.A., Kennelly, S.J., Hodgson, K.E., Ashby, C.J.T., Beatson, M.L., 2001. Retained and discarded catches from commercial beach-seining in Botany Bay, Australia. Fisheries Research 50, 205-219.

- Hastie, T.J., Tibshirani, 1990. *Generalized Additive Models*. Chapman and Hall, London.
- Holst, R., Revill, A., 2009. A simple statistical method for catch comparison studies. *Fisheries Research* 95, 254-259.
- Howell, W.H., Langan, R., 1987. Commercial Trawler Discards of Four Flounder Species in the Gulf of Maine. *North American Journal of Fisheries Management* 7, 6-17.
- Härdle, W., Huet, S., Mammen, E., Sperlich, S., 2004. Bootstrap Inference in Semiparametric Generalized Additive Models. *Econometric Theory* 20, 265-300.
- Härdle, W., Mammen, E., 1993. Testing parametric versus nonparametric regression. *The Annals of Statistics* 21, 1926-1947.
- He I., 2007. Selectivity of large mesh trawl codends in the Gulf of Maine I. Comparison of square and diamond mesh. *Fisheries Research* 83, 44-59.
- Herrmann, B.; Priour, D; Krag, L., 2006. Simulation-based study of the combined effect on cod-end size selection of turning meshes by 90° and reducing the number of meshes in the circumference for round fish. *Fisheries Research* 84, 222-232.
- Pope., 1975. Selectivity of fishing gears. In: FAOs (Ed.), *Manual of methods for fish stock assessment*. FAO, Rome.
- Kaiser, J.M., de Groot, S., 2000. *Effects of Fishing on Target Species and Habitats*. Blachwell Science Ltd, Oxford.
- Keleher, K., 2005. Discards in the World's marine fisheries. An update. FAO Fisheries Technical Paper N°470. FAO, Rome, p. 131.
- Lart, W., R., C., M., A., R., B., Kingston, A., Hugues-Dit-Ciles, E., Lucio, P., N., P., Santurtún, M., and Quincoces, I., 2002. Monitoring of discarding and retention by trawl fisheries in Western Waters and the Irish sea in relation to stock assessment and technical measures.
- Maartens, L., Gamst, K., Schneider, P., 2002. Size selection and release of juvenile monkfish *Lophius vomerinus* using rigid sorting grids. *Fisheries Research* 57, 75-88.

- Machias, A., Vassilopoulou, V., Vatsos, D., Bekas, P., Kallianiotis, A., Papaconstantinou, C., Tsimenides, N., 2001. Bottom trawl discards in the northeastern Mediterranean Sea. *Fisheries Research* 53, 181-195.
- Madsen, N., 2007. Selectivity of fishing gears used in the Baltic Sea cod fishery. *Reviews in Fish Biology and Fisheries* 17, 517-544.
- Massuti, E., Ordines, F., Guijarro, B., 2009. Efficiency of flexible sorting grids to improve size selectivity of the bottom trawl in the Balearic Islands (western Mediterranean), with comparison to a change in mesh codend geometry. *Journal of Applied Ichthyology* 25, 9.
- McCullagh, P., Nelder, J., 1989. *Generalized Linear Models*. Chapman and Hall, London.
- Millar, R.B., 1992. Estimating the Size-selectivity of Fishing Gear by Conditioning on the total catch. *Journal of the American Statistical Association* 87, 962-968.
- Millar, R.B., 1993. Incorporation of between-haul variation using bootstrapping and nonparametric estimation of selection curves. *Fisheries Bulletin* 91, 564-572.
- Millar, R.B., Walsh, S.J., 1992. Analysis of trawl selectivity studies with an application to trouser trawls. *Fisheries Research* 13 91, 205-220.
- Millar, R.B., 1994. Sampling from trawl gears used in sized selectivity experiments. *ICES Journal of Marine Science* 51, 293-298.
- Millar, R.B., Fryer, R.J., 1999. Estimating the size-selection curves of towed gears, traps, nets and hooks. *Reviews in Fish Biology and Fisheries* 9, 89-116.
- Munro, P., Somerton, D., 2001. Maximum likelihood and non-parametric methods for estimating trawl footrope selectivity. *ICES Journal of Marine Science* 58, 220-229.
- Murawski, S.A., 1996. Factors Influencing By-catch and Discard Rates: Analyses from Multispecies/Multifishery Sea Sampling. *Journal of Northwest Atlantic Fishery Science* 19, 31-39.

- Nadaraya, E.A., 1964. On estimating regression. *Theory of Probability and its Applications* 10, 186-190.
- Olaso, I., Velasco, F., Pérez, N., 1998. Importance of discarded Blue Whiting (*Micromesistius poutassou*) in the diet of Lesser-Spotted Dogfish (*Scyliorhinus canicula*) in the Cantabrian sea. *ICES Journal of Marine Science* 55, 331-341.
- O'Neill, F.G., McKay, S.J., Ward, J.N., Strickland, A., Kynoch R.J., Zurur A.F., 2002. An investigation of the relationship between sea state induced vessel motion and cod-end selection. *Fisheries Research* 60, 107-130.
- Ordines, F., Massutti, E., Guijarro, B., Mas, R., 2006. Diamond vs. square mesh codend in a multi-species trawl fishery of the western Mediterranean: effects on catch composition, yield, size selectivity and discards. *Aquatic Living Resources*. 19, 9.
- Pérez, M., Santos, J., Araujo, H., López, V., 2006. Informe final de pesca experimental RAI-AP-10/2006, dirigida a la adecuación del tamaño de malla en la pesquería de gallo mediante arrastre de fondo en NEAFC.
- Pérez, N., Santos, J., Araujo, H., Lustres, M., 2007. Informe final de pesca experimental RAI-AP- 38/2007 Dirigida a la Evaluación del rendimiento y selectividad de nuevas medidas en el arte de arrastre litoral de Cuatro Caras y el arte de arrastre de Raspita en el Caladero Nacional Cantábrico-Noroeste. Instituto Español de Oceanografía.
- Revill, A., Catchpole, T., Dunlin, G., 2007. Recent work to improve the efficacy of square-mesh panels used in a North Sea *Nephrops norvegicus* directed fishery. *Fisheries Research* 85, 321-327.
- Ryer C.H., 2007. A review of flatfish behavior relative to trawls. *Fisheries Research* 90, 138-146.
- Rochet, M., Trenkel, V., 2005. Factors for the variability of discards: assumptions and field evidence. *Canadian Journal of Fisheries and Aquatic Sciences* 62, 224-235.
- Ruppert, D., Wand, M.P., Carroll, R.J., 2003. *Semiparametric regression*. University Press, Cambridge.

- Stratoudakis, Y., Fryer, R., Cook, R., 1998. Discarding practices for commercial gadoids in the North Sea. *Canadian Journal of Fisheries and Aquatic Sciences* 55, 1632-1644.
- Stratoudakis, Y., Fryer, R., Cook, R., Pierce, G., Coull, K., 2001. Fish bycatch and discarding in Nephrops trawlers in the Firth of Clyde (west of Scotland). *Aquatic Living Resources* 14, 283-291.
- Tasker, M.L., Camphuysen, C.J., Cooper, J., Garthe, S., Montevecchi, W.A., Blaber, S.J.M., 2000. The impacts of fishing on marine birds. *ICES Journal of Marine Science* 57, 531-547.
- R Development Core Team., 2009. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.
- Trujillo, V., Perez, N., Pereda, P., 1997. Factors Affecting Discards by the Coastal Spanish North Atlantic Trawlers. ICES CM. ICES, p. 19.
- Wahba, G., 1990. Spline models for observational data. *Regional Conference Series in Applied Mathematics* 59.
- Walsh, S.J., Engås, E., Ferro, R., Fonteyne, R., and Marlen, B.v., 2002. To catch or conserve more fish: the evolution of fishing technology in fisheries science. *ICES Marine Science Symposia*.
- Watson, G.S., 1964. Smooth Regression Analysis. *Sankhya, Series A* 26, 359-372.
- Wood, S., 2006. *Generalized Additive Models*. An introduction with R. Chapman and Hall, London.

# Anexo: Lexislación

Neste apartado expóñense as referencias sobre as regulamentacións pesqueiras citadas ao longo do texto

- 1** Commission Staff Working Document (EC) No. 380. A policy to reduce unwanted by-catches and eliminate discards in European fisheries. Brussels, 28 March 2007.
- 2** Commision Non-Paper (EC). On the implementation of the policy to reduce unwanted by-catch and eliminate discards in European fisheries. Brussels, 25 April 2008.
- 3** ORDEN APA/16/2002, pola que se establece un plan de pesca para a pesqueira de arrastre de fondo no Caladeiro Nacional do Cantábrico e Noroeste.
- 4** Regulamento (CE) n° 2602/2001 da Comisión, de 27 de decembro do 2001, polo cal se establecen medidas técnicas adicionais encamiñadas á recuperación da poboación de merluza nas subzonas CIEM III, IV, V, VI y VII, e nas divisións CIEM VIIIa,b,d,e.
- 5** Regulamento (CE) n° 850/98 da comisión, de 30 de Marzal de 1998, para a conservación de recursos pesqueiros a través de medidas técnicas para a protección de xuvenís de organismos mariños.