



**inbio2.0**

---

**Software Estimación Incertidumbre de Parámetros  
Biológicos**

Paz Sampedro, Valentín Trujillo, María Saínza

# I. Introducción

## ¿Qué es *inbio*? ¿Qué es R?

R es un lenguaje de programación de distribución libre y gratuita, cuyo uso se ha extendido en los últimos años entre investigadores.

*Inbio* es una herramienta para estimar la incertidumbre de los siguientes parámetros biológicos: relación talla-peso, crecimiento en talla y en peso, madurez en talla y en edad y sexratio por talla y edad. Para ello ajusta los datos originales a los modelos más empleados actualmente para poblaciones de especies marinas, y aplica un bootstrap no paramétrico para estimar los coeficientes de variación de los parámetros.

Cuando usas *inbio* también estás usando R. Sin embargo, *inbio* al ser un paquete de rutinas semi-automático reduce al mínimo los conocimientos de R que vas a necesitar.

*inbio* 2.0 ha sido creado usando la versión 2.15.2 de R y bajo Windows.

## ¿Qué no es *inbio*?

No es una herramienta para el análisis exploratorio de datos. Es importante que antes de correr las rutinas se identifiquen posibles outliers, o casos anómalos, que en muchos casos impiden el ajuste o provocan resultados irreales.

*inbio* no entiende de muestreos. Si el tamaño muestral no es el adecuado o los datos no son representativos de la población que se quiere estudiar, puede que los modelos no se ajusten y si se ajustan, que los resultados obtenidos sean erróneos.

## Modificaciones de la versión anterior (*inbio*1.2.1)

Esta nueva versión de *inbio* ha corregido algunos fallos de la versión 1.2.1 y presenta nuevas características:

- se creó con R 2.15.2 y es compatible con arquitecturas de 32 y 64 bits.
- lee los archivos de entrada y escribe los resultados desde el directorio definido por el usuario en cada sesión de R.
- descarga automáticamente la librería *boot*.
- genera los gráficos en formato pdf.
- incluye el cálculo de un coeficiente de variación global para las estimas de sexratio por talla y edad. El coeficiente de variación global se calcula como la media ponderada de los coeficientes de variación (talla/edad) usando como factor de ponderación el número de ejemplares por talla/edad.

- no ajusta la función de suavización en los gráficos de sexratio.
- representa, en los gráficos de sexratio, el porcentaje de hembras (por talla/edad) y el coeficiente de variación asociado.
- presenta nuevas características generales de los gráficos (símbolos, límites de ejes) para su mejor visualización.
- en los casos en los que no es posible el cálculo del coeficiente de variación para el sexratio (talla/edad), debido a que la mediana es cero y la varianza es positiva, se presenta como “NA” (not available), en vez de “inf”.

## II. Guía del Usuario

### Requerimientos previos

- Uso del sistema métrico internacional: decimales con puntos.
- No usar separador de miles.
- La separación de listas es con comas, esto afecta a la creación del archivo de entrada de datos. (Atención: La configuración por defecto en los ordenadores en idioma español es separación de listas por punto y coma).

### Instalación de *inbio*

Descarga el archivo “inbio\_2.0.zip”. Para ello, abre una sesión de R, vete a la barra de menú:

Paquetes

Instalar paquete(s) a partir de archives zip locales...

Y navega por los directorios para localizar el archivo.

### Entrada de Datos

*inbio* lee datos de archivos “.csv”, archivos separados por comas, que se pueden crear desde archivos excel que se graben como “csv”. Los datos de entrada necesarios para correr las cuatro rutinas se estructuran en un único archivo “csv” que deberá estar en el directorio de trabajo definido por el usuario.

Este archivo debe constar de las siguientes columnas, denominadas por las tres primeras letras de la variable a la que representan:

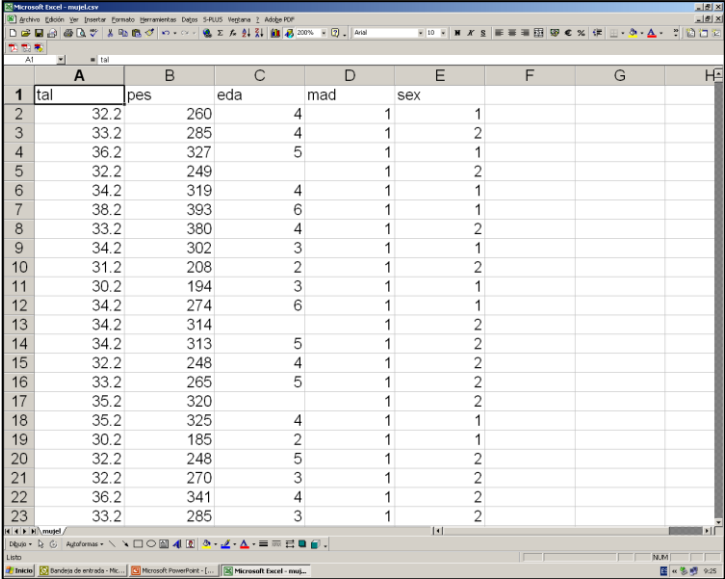
- **tal**: talla en centímetros o milímetros (con o sin decimales)
- **pes**: peso en gramos (con o sin decimales)

- **eda**: edad (0,1,2,...). (Columna opcional, en los casos en que no se tiene datos de edad no se pone)
- **mad**: estado de madurez: 0: inmaduro; 1: maduro
- **sex**: sexo: 1: macho; 2:hembra; 3: indeterminado

Los nombres de las columnas no se pueden cambiar. El orden de las columnas puede ser modificado.

Cada fila corresponde a un ejemplar, y se rellenan las columnas en las que se tengan datos de ese ejemplar. El tratamiento de los missing values es dejar vacía la celda correspondiente (**importante**: no poner ni “-1”, ni “0”).

El nombre del archivo de datos se va utilizar para que las rutinas identifiquen el archivo que deben importar y para nombrar los archivos de salida de la rutina (Posibles nombres son: “merluza.csv”, “sardina03.csv”).



	A	B	C	D	E	F	G	H	I
1	tal	pes	eda	mad	sex				
2		32.2	260	4	1	1			
3		33.2	285	4	1	2			
4		36.2	327	5	1	1			
5		32.2	249		1	2			
6		34.2	319	4	1	1			
7		38.2	393	6	1	1			
8		33.2	380	4	1	2			
9		34.2	302	3	1	1			
10		31.2	208	2	1	2			
11		30.2	194	3	1	1			
12		34.2	274	6	1	1			
13		34.2	314		1	2			
14		34.2	313	5	1	2			
15		32.2	248	4	1	2			
16		33.2	265	5	1	2			
17		35.2	320		1	2			
18		35.2	325	4	1	1			
19		30.2	185	2	1	1			
20		32.2	248	5	1	2			
21		32.2	270	3	1	2			
22		36.2	341	4	1	2			
23		33.2	285	3	1	2			

**Importante:** Una vez hayas creado el archivo comprueba que los datos se encuentran separados por comas, si no *inbio* no es capaz de leerlo.

## Una Sesión de *inbio*

- Abre una sesión de R.
- Define el directorio de trabajo:  
A través de Menús:  
Archivo\Cambiar dir

Usando el comando `setwd()`

Por ejemplo: `> setwd("C:/Users/psampedro/Documents/inbio")`

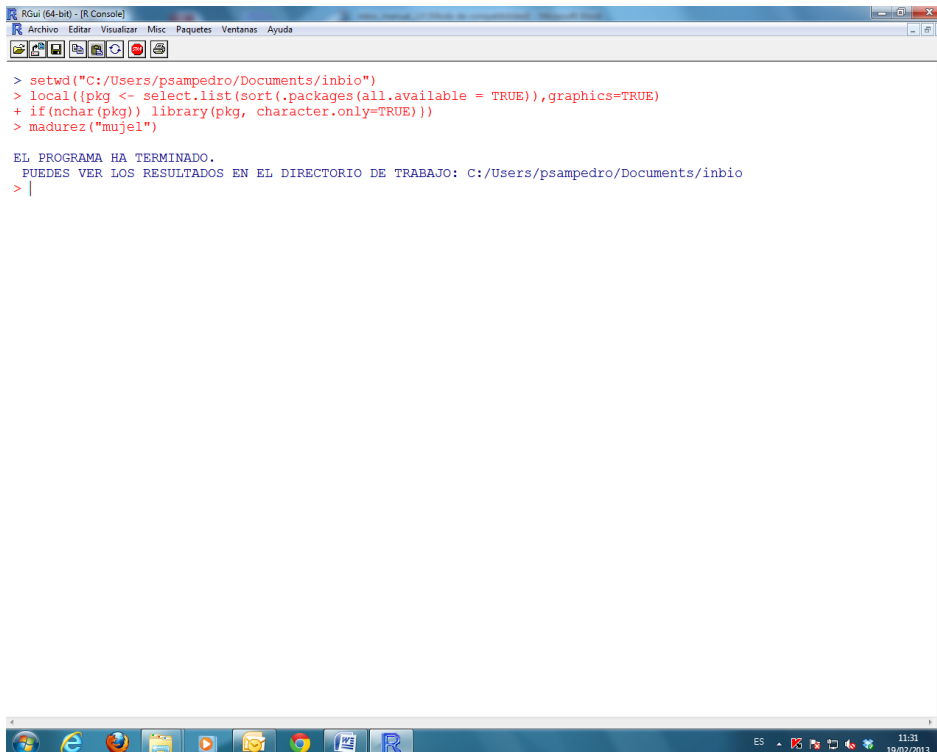
- Descarga *inbio*, para ello teclea:

```
> library(inbio)
```

- Llama las rutinas de *inbio*, modificando los argumentos necesarios en cada caso (ver explicación detallada en el apartado III de este manual). Aquellos argumentos en los que estés de acuerdo con su valor por defecto no hace falta escribirlos:

# Si tu archivo de datos es "C:/Users/psampedro/Documents/inbio/mujel.csv"

```
> madurez(especie="mujel")
```



```
RGui (64-bit) - R Console
> setwd("C:/Users/psampedro/Documents/inbio")
> local({pkg <- select.list(sort(.packages(all.available = TRUE)),graphics=TRUE)
+ if(nchar(pkg)) library(pkg, character.only=TRUE)})
> madurez("mujel")

EL PROGRAMA HA TERMINADO.
PUEDES VER LOS RESULTADOS EN EL DIRECTORIO DE TRABAJO: C:/Users/psampedro/Documents/inbio
> |
```

- Cuando el proceso que has llamado termina, aparece un mensaje en la pantalla indicando que ha finalizado y en qué directorio guardó los archivos con los resultados. El tiempo que tarda el proceso depende de la memoria RAM de tu ordenador, del volumen de datos y de la rutina (p.e.: la rutina de madurez es la que más tarda), varía entre segundos a algunos minutos.

- Ahora puedes llamar otra rutina o cambiar de archivo de datos. Ten en cuenta que si corres dos veces una rutina para el mismo archivo de datos el programa escribe sobre los primeros resultados.

- Para finalizar la sesión, cierra el programa desde la barra de menú: File, exit. Aparecerá el siguiente mensaje: "Save workspace image?". Haz clic en "no", *inbio* no crea ningún objeto de R que tenga que ser guardado, todas las salidas son en archivos "txt" y "pdf".

### III. Rutinas de *inbio*

El package *inbio* contiene cuatro rutinas:

- peso
- madurez
- crecimiento
- sexratio

#### Rutina peso

##### *Descripción*

La rutina peso calcula la relación talla-peso por estimación no lineal y ajuste de mínimos cuadrados (algoritmo de Gauss-Newton). Mediante bootstrap no paramétrico estima el coeficiente de variación de los parámetros de la relación talla-peso.

##### *Uso*

```
> peso(especie = "nombre especie", cl = 1, unid="cm", a=5e-4, b=3, sex = F, n = 1000)
```

##### *Argumentos*

- **especie**: nombre del archivo donde están los datos de entrada, sin la extensión y sin el directorio.
- **cl**: intervalo de talla de la especie, normalmente 0.5 ó 1. Por defecto es 1. Los datos se agrupan en función de este intervalo por lo que los resultados para un mismo archivo de datos difiere si se usa uno u otro intervalo.
- **unid**: unidad de los datos de talla. Por defecto es "cm".
- **a**: Valor de partida ("start value" o "semilla") de la ordenada que usa el proceso para iniciar el ajuste.
- **b**: Valor de partida ("start value" o "semilla") de la pendiente que usa el proceso para iniciar el ajuste.
- **sex**: si sex=F, el proceso analiza los datos de machos y hembras conjuntamente. Si sex=T, analiza separadamente machos y hembras. Por defecto sex=F.
- **n**: número de réplicas bootstrap. El número recomendado es 1000.

## *Detalles*

Es importante que los valores de partida ("start value" o "semilla") no sean muy diferentes de los que se quiere calcular (se pueden consultar valores de otros años o zonas), si no, el proceso iterativo por el que se realiza el ajuste puede fallar.

## *Resultados*

Archivo "txt" en el directorio de trabajo, con los resultados de los datos originales y de las réplicas bootstrap y su coeficiente de variación.

Un archivo "pdf" en el directorio de trabajo, con las gráficas del ajuste con datos originales y bootstrap, las gráficas del análisis de residuos del ajuste de los datos originales y análisis de las réplicas bootstrap.

Si el proceso se realiza separadamente para machos y hembras, los archivos de salida también son independientes para cada sexo.

## *Ejemplos*

```
#Requiere que exista el archivo "merluza.csv" en el directorio de
#trabajo. Calcula la relación talla-peso separadamente para machos
#y hembras
> peso(especie="merluza", sex=T)
```

```
#Requiere que exista el archivo "choco.csv" en el directorio de
#trabajo. Calcula la relación talla-peso para una especie con
intervalo de talla de 0.5 cm y dando valores de partida
orientativos.
```

```
> peso(especie="choco", cl=0.5, a=0.04, b=2.4)
```

## **Rutina madurez**

### *Descripción*

Esta rutina calcula la talla y la edad de madurez sexual aplicando un glm con errores binomiales (regresión logística). Mediante bootstrap no paramétrico estima los coeficientes de variación de: los parámetros de las funciones, de la talla (L50) y la edad (A50) a la que el 50% son maduros.

### *Uso*

```
> madurez(especie = "nombre especie", cl = 1, unid="cm", edad=T, sex = F, n = 1000)
```

### *Argumentos*

- **especie:** nombre del archivo donde están los datos de entrada, sin la extensión y sin el directorio.
- **cl:** intervalo de talla de la especie, normalmente 0.5 ó 1. Por defecto es 1. Los datos se agrupan en función de este intervalo por lo que los resultados para un mismo archivo de datos difiere si se usa uno u otro intervalo.
- **unid:** unidad de los datos de talla. Por defecto es "cm".
- **edad:** si edad=T, requiere datos de edad, calcula la edad de madurez. Si edad=F, no calcula la edad de madurez. Por defecto edad = T.
- **sex:** si sex=F, el proceso analiza los datos de machos y hembras conjuntamente. Si sex=T, analiza separadamente machos y hembras. Por defecto sex=F.
- **n:** número de réplicas bootstrap. El número recomendado es 1000.

### *Resultados*

Dos archivos "txt" en el directorio de trabajo, con los resultados de los datos originales y de las réplicas bootstrap y su coeficiente de variación. Un archivo contiene los resultados de la madurez en talla y otro los de la madurez en edad.

Dos archivos "pdf" en el directorio de trabajo, con las gráficas del ajuste con datos originales y bootstrap, las gráficas del análisis de residuos del ajuste de los datos originales y análisis de las réplicas bootstrap, un archivo contiene las gráficas para el análisis de la madurez en talla y otro las gráficas para madurez en edad.

Si el proceso se realiza separadamente para machos y hembras, los archivos de salida también son independientes.

### *Ejemplos*

```
#Requiere que exista el archivo "lepwhi2004.csv" en el directorio  
#de trabajo. Calcula la talla y la edad de madurez sexual y su  
#coeficiente de variación conjuntamente para ambos sexos.  
> madurez(especie="lepwhi2004")
```

```
#Calcula sólo la talla madurez sexual y su coeficiente de  
#variación, separadamente para machos y hembras.  
> madurez(especie="lepwhi2004", edad=F, sex=T)
```



## Rutina crecimiento

### Descripción

Esta rutina calcula el crecimiento en talla y el crecimiento en peso ajustando la ecuación de von Bertalanffy, por estimación no-lineal y mínimos cuadrados (algoritmo de Gauss-Newton).

Para el crecimiento en talla la ecuación de von Bertalanffy empleada es:

$$L_t = L_{inf} (1 - e^{-k(t - t_0)})$$

donde:

$L_{inf}$  = talla máxima teórica (longitud infinita);

$K$  = coeficiente de crecimiento;

$t$  = edad en años;

$t_0$  = edad en años para una talla teórica: 0;

Para el crecimiento en peso la ecuación de von Bertalanffy empleada es:

$$W_t = W_{inf} (1 - e^{-k(t - t_0)})^b$$

donde:

$W_{inf}$  = peso máximo teórico (peso infinito);

$K$  = coeficiente de crecimiento;

$t$  = edad en años;

$t_0$  = edad en años para un peso teórico: 0;

$b$  = es una constante, la pendiente de la relación talla-peso. Por defecto se emplea  $b=3$ , pero se puede cambiar en los argumentos de la rutina.

### Uso

```
> crecimiento(especie = "nombre especie", cl = 1, unid="cm", sex =  
F, b = 3, Li = 80, Lfijs=F, Ki = 0.7, T0i = -1, Wi = 2500,  
Wfijs=F, Kwi = 0.4, T0wi = -1, n = 1000)
```

### Argumentos

- **especie**: nombre del archivo donde están los datos de entrada, sin la extensión y sin el directorio.
- **cl**: intervalo de talla de la especie, normalmente 0.5 ó 1. Por defecto es 1. Los datos se agrupan en función de este intervalo por lo que los resultados para un mismo archivo de datos difiere si se usa uno u otro intervalo.
- **unid**: unidad de los datos de talla. Por defecto es "cm".
- **sex**: si sex=F, el proceso analiza los datos de machos y hembras conjuntamente. Si sex=T, analiza separadamente machos y hembras. Por defecto sex=F.

- **b**: parámetro de la relación talla-peso empleado en el crecimiento en peso. Por defecto  $b=3$ .
- **Li**: valor de partida ("start value" o "semilla") de la Linf (longitud asintótica) que usa el proceso para iniciar el ajuste del crecimiento en talla.
- **Lfija**: si  $Lfija=T$ , fija el valor dado en "Li" como Linf en la ecuación de von Bertalanffy de crecimiento en talla, usándolo como una constante. Si  $Lfija=F$ , calcula la Linf y su coeficiente de variación. Por defecto  $Lfija=F$ .
- **Ki**: valor de partida ("start value" o "semilla") de la Ki (coeficiente decrecimiento) que usa el proceso para iniciar el ajuste del crecimiento en talla.
- **T0i**: valor de partida ("start value" o "semilla") de la T0i (punto en el tiempo, en el que el pez tiene una talla cero) que usa el proceso para iniciar el ajuste del crecimiento en talla.
- **Wi**: valor de partida de ("start value" o "semilla") del Winf (peso asintótico) que usa el proceso para iniciar el ajuste del crecimiento en peso.
- **Wfijo**: si  $Wfijo=T$ , fija el valor dado en "Wi" como Winf en la ecuación de von Bertalanffy de crecimiento en peso, usándolo como una constante. Si  $Wfijo=F$ , calcula el Winf y su coeficiente de variación. Por defecto  $Wfijo=F$ .
- **Kwi**: valor de partida de ("start value" o "semilla") del Kwi (coeficiente de crecimiento) que usa el proceso para iniciar el ajuste del crecimiento en peso.
- **T0wi**: valor de partida de ("start value" o "semilla") del T0i (punto en el tiempo, en el que el pez tiene un peso cero) que usa el proceso para iniciar el ajuste del crecimiento en peso.
- **n**: número de réplicas bootstrap. El número recomendado es 1000.

### *Detalles*

Es importante que los valores de partida ("start value" o "semilla") no sean muy diferentes de los que se quiere calcular (se pueden consultar valores de otros años o zonas), si no, el proceso iterativo por el que se realiza el ajuste puede fallar.

El ajuste de la ecuación de crecimiento en peso de von Bertalanffy presenta problemas en muchos casos (ver explicación más detallada en Quinn & Deriso (1999)).

### *Resultados*

Dos archivos "txt" en el directorio de trabajo, con los resultados de los datos originales y de las réplicas bootstrap y su coeficiente de variación, un archivo para los resultados del crecimiento en talla y otro para los resultados del crecimiento en peso.

Dos archivos "pdf" en el directorio de trabajo con las gráficas del ajuste con datos originales y bootstrap, las gráficas del análisis de residuos del ajuste de los datos

originales y análisis de las réplicas bootstrap. Un archivo con las gráficas del crecimiento en talla y otro archivo con las gráficas del crecimiento en peso.

Si el proceso se realiza separadamente para machos y hembras, los archivos de salida también son independientes.

### *Ejemplos*

```
#Requiere que exista el archivo "bosci2004.csv".Calcula los
#parámetros de la ecuación de crecimiento en talla y peso de von
#Bertalanffy y su coeficiente de variación #conjuntamente para
#ambos sexos.
```

```
> crecimiento(especie="bosci2004", Li = 40, Ki = 0.27, T0i = -1.9,
Wi = 2500, Kwi = 0.4, T0wi = -1)
```

```
#Calcula los parámetros de la ecuación de crecimiento en talla y
#peso de von Bertalanffy y su coeficiente de variación
#independientemente para machos y hembras. Fija en la ecuación de
#crecimiento en talla la Linf en 150 cm y en la ecuación de
#crecimiento en peso el Winf en 25000 g.
```

```
> crecimiento(especie="congrío", sex=T, Li = 150, Lfija=T, Ki =
0.27, T0i = -1.9, Wi = 25000, Wfijo=T, Kwi = 0.4, T0wi = -1)
```

## **Rutina sexratio**

### *Descripción*

Calcula el porcentaje de hembras por clase talla y por edad, sus coeficientes de variación mediante bootstrap no paramétrico, y los valores globales por talla y edad. Para estos cálculos se considera el sexratio como el porcentaje de hembras sobre el total de machos y hembras (**Nota: los indeterminados no se tienen en cuenta en el cálculo**)

### *Uso*

```
> sexratio(especie = "nombre especie", cl = 1, unid="cm", edad=T,
n = 1000)
```

### *Argumentos*

- **especie:** nombre del archivo donde están los datos de entrada, sin la extensión y sin el directorio.
- **cl:** Intervalo de talla usado para el sexratio por talla. Por defecto es 1.
- **unid:** unidad de los datos de talla. Por defecto es "cm".
- **edad:** si edad=T, requiere datos de edad, calcula el sexratio por edad. Si edad=F, no calcula el sexratio por edad. Por defecto edad = T.

- **n**: número de réplicas bootstrap. El número recomendado es 1000.

### Resultados

Archivo "txt" en el directorio de trabajo, con dos tablas: una tabla con los resultados de sexratio por talla y otra con los resultados de sexratio por edad.

Archivo "pdf" en el directorio de trabajo con dos gráficas. Se representa el sexratio por talla/edad estimado para los datos originales y su correspondiente coeficiente de variación.

### Ejemplos

```
#Requiere que exista el archivo "merluza.csv". Calcula el sexratio  
#por intervalo de talla de 1 cm y el sexratio por edad.  
> sexratio(especie="merluza")
```

```
#Requiere que exista el archivo "choco.csv". Calcula el sexratio  
#por intervalo de talla de 0.5 cm. No calcula el sexratio por  
#edad.  
> sexratio(especie="choco", cl=0.5, edad=F)
```

## IV. Bibliografía

### R

Venables, B., Smith, D., Gentleman, R. and R. Ihaka. *Notes on R: A Programming Environment for Data Analysis and Graphics*. Dept. of Statistics, University of Adelaide and University of Auckland, 1997. Disponible en <http://cran.at.r-project.org/doc/R-intro.pdf>.

W.N. Venables and B.D. Ripley. 'R' complements to Modern Applied Statistics with S-PLUS. En <http://www.stats.ox.ac.uk/pub/MASS3>, 1999.

### inbio

Sampedro P., Saínza M., Trujillo V. 2005. A simple tool to calculate biological parameter's uncertainty. Working Document in Workshop on Sampling Design for Fisheries Data. ICES CM 2005/ACFM:11, Pasajes (Spain), 82 pp.

### Incertidumbre de parámetros biológicos

Brodziak, J. & R. Mikus. 2000. Variation in life history parameters of Dove sole, *Microstomus pacificus*, off the coast of Washington, Oregon, and northern California. *Fishery Bulletin* 98: 661-673.

Bullock, S. H., Turner, R.M., Hastings, J.R., Escoto-Rodríguez, M., Ramírez Apud López, Z. & J.L. Rodríguez Navarro. 2004. Variance on size-age curves: Bootstrapping with autocorrelation. *Ecology*, 85(8): 2114-2117.

Davison, A.C. & D.V. Hinkley. 1999. Bootstrap methods and their application. Cambridge University Press, 575 pp.

Efron, B. 1979. Bootstrap methods: another look at the jackknife. *Annals of Statistics*, 7. 1-26 p.

Gros, P., Laurec, A. & H. Dupouy. 1987. Estimation des paramètres de croissance apport des techniques de reechantillonnage (Bootstrap & Jackknife) et des statistiques robustes (Médianes et modes). *ICES Comité des statistiques*, CM 1987/D: 10, 25pp.

Meyer, J.S., Ingersoll, C.G., McDonald, L.L. & M.S. Boyce. 1986. Estimating uncertainty in population growth rates: Jackknife vs Bootstrap techniques. *Ecology*, 67(5): 1156-1166.

Moguedet, P. & N. Pérez. Estimation of Megrim (*Lepidorhombus whiffiagonis*) growth parameters, for males and females, from the ICES division VII: fitting to the V. Bertalanffy model using, resampling techniques, as well as several adjustable central values (mean, median and mode length at age). ICES CM 1988/G: 9. 18 pp.

Quinn, T.J. & R.B. Deriso. 1999. Quantitative Fish Dynamics. Oxford University Press, 542 pp.

Roa, R., Ernst, B & F. Tapia. 1999. Estimation of size at sexual maturity: an evaluation of analytical and resampling procedures. *Fishery Bulletin* 97: 570-580.

## Anexo 1

Día: 18-feb-2013 Hora: 14:22:52

Día y hora de finalización del análisis (datos del sistema operativo del ordenador).

EDAD DE MADUREZ

-----  
Especie: mujel  
Sexo: Ambos  
N: 3644

Número de ejemplares válidos empleados en el análisis.

Argumentos Rutina:  
madurez(especie = "mujel")

Valor adjudicado a los argumentos en el análisis. Los argumentos que no aparecen mantienen su valor por defecto.

RESULTADOS AJUSTE(datos originales)

Call:

```
glm(formula = mad ~ eda, family = binomial(link = logit), data =
mad.eda.dat)
```

Deviance Residuals:

Min	1Q	Median	3Q	Max
-3.6225	-0.4040	0.0216	0.3194	2.2568

Coefficients:

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-3.36738	0.13354	-25.22	<2e-16 ***
eda	1.80496	0.07276	24.81	<2e-16 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 4770.3 on 3643 degrees of freedom  
Residual deviance: 2285.3 on 3642 degrees of freedom  
AIC: 2289.3

Number of Fisher Scoring iterations: 7

PARÁMETROS ESTIMADOS: ORIGINALES-BOOTSTRAP

	Estima datos	Estima boot	CV boot
B0	-3.3674	-3.3714	0.046
B1	1.805	1.8039	0.054
E50 (año)	1.8656	1.8646	0.017