

UNIVERSIDAD DE LA REPÚBLICA

Facultad de Ciencias y Facultad de Ciencias Económicas y

Administración

Instituto de Estadística

Informe de Pasantía:

**Ajustes de Modelos Dinámicos
aplicados a recursos pesqueros**

Institución donde se realizó la pasantía:

Dirección Nacional de Recursos Acuáticos

Ministerio de Ganadería Agricultura y Pesca

Tutores:

María Inés Lorenzo - Marco Scavino

Pasantes:

Fernando Brito - María Saravia

2012

Resumen

El objetivo principal del presente trabajo es lograr el ajuste de modelos dinámicos de Biomasa de Schaefer (1954) aplicados a recursos pesqueros. A tal efecto se han utilizado datos de la literatura y datos referidos a la especie *Cynoscion guatucupa* (pescadilla) comercializada en el Uruguay. En primer lugar se ha estudiado el comportamiento de las funciones de -log-verosimilitud correspondiente a dos versiones estocásticas del modelo de Schaefer, una con incertidumbre en las observaciones y otra con incertidumbre en el proceso. En cada caso se han identificado regiones óptimas de verosimilitud para los parámetros de interés, la capacidad de carga y la tasa intrínseca de crecimiento.

La segunda parte del trabajo ha sido enfocada al análisis Bayesiano del modelo dinámico de Biomasa de Schaefer con incertidumbre en las observaciones. Dicho análisis se ha implementado mediante el programa OpenBUGS que, en este caso ofrece ventajas computacionales respecto al entorno de programación R utilizado en el resto del trabajo.

Los resultados obtenidos sugieren que las técnicas estadísticas aplicadas en este trabajo permiten encontrar los puntos de referencia biológicos adecuados para la pesquería de la especie de interés, tal como el Rendimiento Máximo Sostenible, índice de suma importancia para la preservación de una pesquería.

Palabras clave: Modelo de Crecimiento Logístico, Modelo de Schaefer, ecuaciones diferenciales, ecuaciones en diferencias, verosimilitud, perfiles de verosimilitud, intervalos de confianza, test de normalidad, incertidumbre, teorema de Bayes, método MCMC, programación en R y OpengBUGS.

Índice general

1. Introducción	6
2. Marco Teórico	8
2.1. Modelos matemáticos dinámicos de poblaciones pesqueras	8
2.1.1. Supuestos biológicos para modelos dinámicos de producción	9
2.2. Modelo de crecimiento logístico	9
2.2.1. Ecuación logística	11
2.2.2. Ecuación en diferencias de crecimiento logístico con explotación . .	16
2.3. Modelo de Schaefer	17
2.4. Enfoque Bayesiano	23
2.4.1. Elementos de la Teoría Bayesiana	23
2.4.2. Muestreo de Gibbs	26
2.4.3. Método MCMC y BUGS	26
3. Antecedentes	30
4. Metodología	31

4.1. Base de datos de la literatura	31
4.2. Base de datos de la Pesquería Uruguaya	32
4.3. Descripción del procedimiento para ajustar el modelo usando Máxima Verosimilitud	34
4.4. Programación en R	37
4.5. Programación en OpenBUGS	38
5. Resultados	39
5.1. Aplicación con datos de la literatura	39
5.1.1. Unidades	39
5.1.2. Ajuste del modelo de Schaefer por Máxima Verosimilitud con incertidumbre en las observaciones	39
5.1.3. Ajuste del modelo de Schaefer por Máxima Verosimilitud con incertidumbre en el proceso	46
5.1.4. Ajuste del modelo con el enfoque Bayesiano	53
5.2. Aplicación con datos de la Pesquería Uruguaya	65
5.2.1. Unidades	65
5.2.2. Ajuste del modelo de Schaefer por Máxima Verosimilitud con incertidumbre en las observaciones	65
5.2.3. Ajuste del modelo de Schaefer por Máxima Verosimilitud con incertidumbre en el proceso	76
5.2.4. Ajuste del modelo con el enfoque Bayesiano	76
6. Conclusiones	87

7. Sugerencias para trabajos futuros	89
8. Anexo	92
8.1. Script de R para realizar el ajuste por máxima verosimilitud	92
8.1.1. Funciones	92
8.1.2. Datos de la literatura	115
8.1.3. Script principal para ajuste con incertidumbre en la observaciones .	116
8.1.4. Script principal para ajuste con incertidumbre en el proceso	121
8.2. Script de R para estimar parámetros de distribuciones a priori	126
8.3. Script de OpenBUGS para ajustar el modelo con el enfoque Bayesiano . .	128

Glosario ¹

- **Biomasa (B):** Peso o cantidad de un individuo o de un grupo de individuos contemplados en un stock, usualmente se mide en unidades de masa como toneladas.
- **Capacidad de carga (K):** Situación de equilibrio entre la reproducción potencial del stock y la capacidad del ecosistema para mantenerlo. Teóricamente es el límite de biomasa del stock en ausencia de mortalidad pesquera.
- **Captura (C):** Cantidad de individuos capturados, usualmente se mide en unidades de masa como toneladas.
- **Captura Por Unidad de Esfuerzo ($CPUE$):** Es un índice de abundancia, que está relacionado con la biomasa disponible del recurso, preferentemente proporcional a ella.
- **Coefficiente de capturabilidad (q):** Fracción de la biomasa capturada por unidad de esfuerzo de pesca.
- **Punto de referencia biológico:** es un determinado valor de interés biológico, por ejemplo la biomasa utilizada para la gestión de una pesquería, tomando en consideración la mejor captura posible y/o asegurando la conservación del recurso pesquero.
- **Stock:** Conjunto de supervivientes de las cohortes de un recurso pesquero, en un cierto instante o período de tiempo. Puede referirse a la biomasa o a un número de individuos.

¹Fuente: [4], Páginas. xi a xiii

- **Tasa intrínseca de crecimiento de la biomasa o población (r):** Es una constante de los modelos de producción, es la diferencia entre las tasas de nacimiento y de muerte. Por ejemplo determina la relación de crecimiento de una cohorte.
- **Tasa de nacimiento (b):** Es la tasa relativa a la variación del número de supervivientes que nacen.
- **Tasas de muerte (d):** Es la tasa relativa a la variación del número de supervivientes que mueren por diferentes razones.

Capítulo 1

Introducción

En el presente trabajo se implementa un tipo de modelo para la evaluación del stock de recursos pesqueros que pertenece a la clase de los denominados “modelos de producción excedente”. Estos modelos son una herramienta tradicional y sencilla para la evaluación del stock de un recurso ecológico, esto se debe a que para su implementación sólo se necesita una serie de capturas y una serie de índices de abundancia. En particular se profundizará en un modelo basado en el crecimiento logístico de una población sujeta a explotación, originalmente estudiado por Milner Baily Schaefer en 1954. El objetivo perseguido en la aplicación de este tipo de modelos es determinar el nivel óptimo de esfuerzo, es decir, el esfuerzo que produce el máximo rendimiento que puede ser sostenible sin afectar la productividad a largo plazo del stock del recurso, lo que se denomina Rendimiento Máximo Sostenible *RMS* ([4] Cap. 5).

En el manejo pesquero resulta de interés prioritario la determinación del *RMS*, porque permite una utilización adecuada del recurso de forma que se pueda explotar el mismo a través del tiempo en forma sostenida, sin agotar el recurso. Tal conocimiento permite un adecuado dimensionamiento de la flota que tiene a esa especie como objetivo, y resulta útil para establecer cupos de captura totales y por país cuando se trata de un recurso compartido. Es por ello que en el ámbito de la Comisión Técnico Mixta del Frente Marítimo (Comisión binacional creada en el ámbito del Tratado del Río de la Plata y su

Frente Marítimo, que suscribieran los Gobiernos de la República Argentina y la República Oriental del Uruguay el 19 de noviembre de 1973) se considera de sumo interés establecer estos indicadores, a efectos de un manejo integrado de la especie [1].

El procesamiento de la información comienza con las series de tiempo de las variables *Captura por Unidad de Esfuerzo* y *Captura*, y a través de procedimientos clásicos implementados mediante el lenguaje R, se obtienen las estimaciones de los parámetros de interés del Modelo de Schaefer que son la *tasa intrínseca de crecimiento* r y la *capacidad de carga* K . En el enfoque clásico se obtienen los respectivos perfiles de verosimilitud y se construyen intervalos de confianza para los parámetros óptimos encontrados. Para visualizar la región donde se encuentran los óptimos se han construido gráficas de superficies de nivel. Los parámetros estimados permiten determinar puntos de referencia biológicos, entre los más importantes se encuentra RMS , estos valores son de especial interés para los investigadores de la DINARA.

Por otro lado se realiza un análisis Bayesiano para encontrar las distribuciones a posteriori de los parámetros r , K , RMS , obteniéndose también medidas de tendencia central para estos. El análisis Bayesiano se realiza utilizando el lenguaje de programación de OpenBUGS.

Capítulo 2

Marco Teórico

2.1. Modelos matemáticos dinámicos de poblaciones pesqueras

Los modelos matemáticos utilizados para describir e interpretar la dinámica de poblaciones biológicas, son sólo una representación abstracta de la población de interés. Para la implementación de los mismos se necesita realizar suposiciones básicas que simplifican la realidad, las cuales no pueden ser contradictorias y no pueden ser demostradas ([4] Pág. 6 y [8] Cap. 1).

Varias clases de modelos se utilizan para la evaluación pesquera, en la presente investigación se adopta un modelo perteneciente a los denominados *modelos de producción*. Este modelo se propone estudiar la evolución de la Biomasa total del stock en su globalidad, contemplando entre otros efectos, el esfuerzo de pesca. Dicho modelo no considera la estructura de edades o de tallas del stock, las cuales dan lugar a la aplicación de los denominados *modelos estructurados*.

2.1.1. Supuestos biológicos para modelos dinámicos de producción

- Todos los individuos en la población tienen ecológicamente las mismas características, se trata de una única población que no distingue individuos por peso, morfología, entre otras. Tampoco se tienen en cuenta efectos migratorios.
- No hay tiempos sustanciales de retraso en los procesos poblacionales, es decir toda la población responde inmediatamente a cualquier cambio independientemente de la estación y de la ubicación geográfica.
- Los parámetros del modelo son constantes, estos ignoran el factor estacional y ambiental.

2.2. Modelo de crecimiento logístico

Uno de los primeros modelos matemáticos sobre crecimiento de la población fue planteado por Thomas Robert Malthus en 1798 (ver, por ejemplo [2], Cap. 5), este modelo de crecimiento geométrico de poblaciones fue tomado como punto de partida para los posteriores modelos de crecimiento de poblaciones.

El supuesto más simple de crecimiento de una población es asumir que la misma crece ilimitadamente. Este tipo de poblaciones puede describirse mediante un modelo exponencial. Si las tasas de nacimiento y muerte son constantes, esto es, la población crece a un ritmo constante en cada período, la tasa de crecimiento de la población puede expresarse mediante la siguiente ecuación diferencial:

$$\frac{dB(t)}{dt} = (b - d)B(t) = rB(t) \quad (2.1)$$

donde $B(t)$ es la Biomasa, que es la cantidad de individuos en un determinado instante de tiempo, habitualmente se mide en unidades de masa, por ejemplo toneladas, b y d son la tasa de nacimiento y tasa de muerte respectivamente y r es la tasa intrínseca (o neta) de

crecimiento de la población, esta tasa puede no ser constante sobre un período de tiempo extenso ya que varía con el medio ambiente, la comida disponible, etc.

Habitualmente la notación para indicar la cantidad de individuos en una población en un instante t es $N(t)$, a los efectos del presente trabajo se utiliza $B(t)$.

La ecuación 2.1 es una ecuación diferencial homogénea de primer orden. La solución de la ecuación diferencial 2.1, que describe la Biomasa de una población en cualquier período con la condición inicial de $B_{t_0} = B_0$ es:

$$\begin{cases} B_{t_0} = B_0 \\ B(t) = B_0 \exp((b - d)t) = B_0 \exp(rt) \end{cases} \quad (2.2)$$

Si se multiplica y divide la ecuación 2.2 por $\exp(r)$ se obtiene,

$$B_0 \exp(r(t - 1)) \exp(r) = B_{t-1} \exp(r)$$

$$B(t) = B_{t-1} \exp(r)$$

Es por esta expresión 2.2, que el modelo se denomina *Modelo de crecimiento exponencial*. Del mismo se desprende que dependiendo de la relación entre la tasa de nacimiento y la tasa de muerte, el crecimiento exponencial puede describirse como: una población que se extingue, una población que se mantiene estable o una población que crece.

El modelo de crecimiento logístico, a diferencia del modelo de crecimiento exponencial, se adecúa mayormente a la realidad para estimar el tamaño poblacional. Según su formulación, para remover el supuesto de crecimiento ilimitado de la población, Pierre-François Verhulst publicó en 1838 la obra "Note on the law of population growth" ([2], Cap. 6), en la cual introduce un umbral de saturación, es decir pone un "*tope*" a la población, cuyo fundamento se basa en factores tales como: obstáculos de supervivencia,

restricciones ambientales, entre otros, que actúan como limitante en dicho crecimiento. Es por estas limitaciones que se denomina *Modelo de crecimiento denso-dependiente* ([8] Pág. 25 a 36 y [10] Pág. 749 a 754), es decir a menor densidad de población mayor es el crecimiento y a mayor densidad el crecimiento de la población es menor. La formulación general de un modelo de crecimiento poblacional denso - dependiente es la siguiente:

- *En tiempo discreto*

$$B_{t+1} = [f(B_t)]B_t \quad (2.3)$$

- *En tiempo continuo*

$$\frac{dB(t)}{dt} = [g(B(t - T))]B(t) \quad (2.4)$$

con T, tiempo de resago.

donde f y g son las funciones de denso - dependencia (para más detalle ver [10]).

2.2.1. Ecuación logística

La formulación general del modelo de crecimiento poblacional denso-dependiente originalmente planteada por Verhulst ([2], Cap. 6) tiene la siguiente forma:

$$\frac{dP(t)}{dt} = rP(t) \left(1 - \frac{P(t)}{K}\right)$$

donde $P(t)$ es la población que depende del tiempo t . Si se generaliza esta expresión considerando a $P(t)$ como la Biomasa $B(t)$, se puede plantear la siguiente ecuación:

$$\begin{cases} \frac{dB(t)}{dt} = r(t)B(t) \left(1 - \frac{B(t)}{K(t)}\right) \\ B(t_0) = B_0 \end{cases} \quad (2.5)$$

donde K es la capacidad de carga (nivel de saturación del sistema) y además r y K varían en el tiempo, en este caso la función de denso - dependencia es el término $r(t) \left(\frac{K(t)-B(t)}{K(t)}\right)$. Para más detalle ver [3] Cap. 1 y 2. Reescribiendo la ecuación 2.5 queda de la siguiente

forma:

$$\begin{cases} \frac{dB(t)}{dt} = r(t)B(t)\left(1 - \frac{B(t)}{K(t)}\right) \\ B(t_0) = B_0 \end{cases} \quad (2.6)$$

La aplicación que se suele utilizar de la ecuación 2.6 es asumir r y K constantes en el tiempo, se observa que para $B(t) = K$ y $B(t) = 0$ la tasa de crecimiento de la población $\frac{dB(t)}{dt} = 0$ y además la derivada de segundo orden $\frac{d^2B(t)}{d^2t} > 0$ para $B < K/2$ y $\frac{d^2B(t)}{d^2t} < 0$ cuando $B(t) > K/2$, el punto máximo de crecimiento se da en $K/2$ es decir cuando $\frac{d^2B(t)}{d^2t} = 0$. Este punto máximo es el que se indica en la Figura 2.1 cuando el tamaño de la población toma el valor $K/2$.

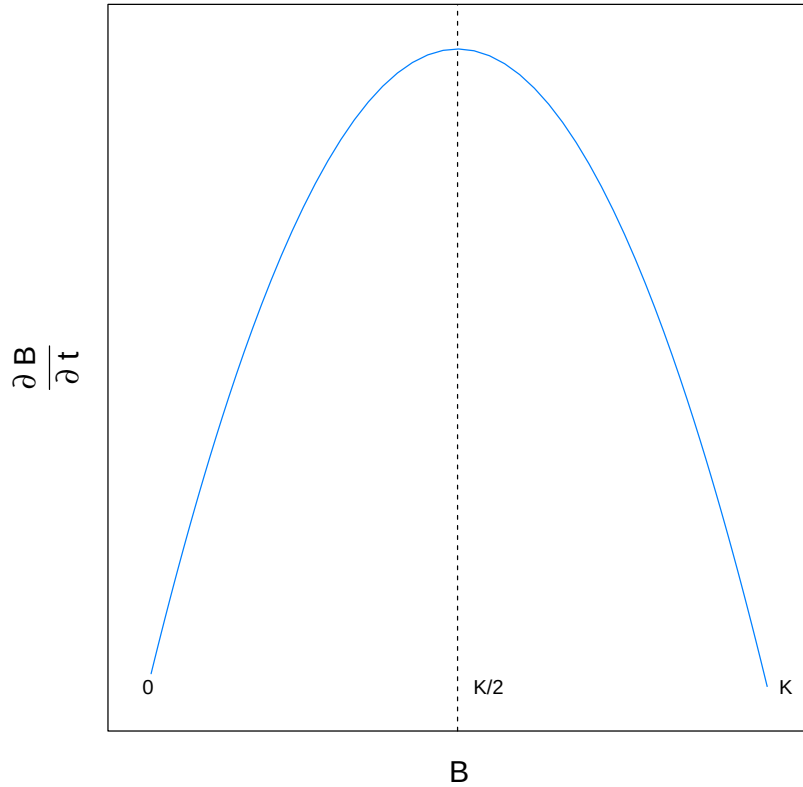


Figura 2.1: Ecuación 2.6 para el caso particular de r y K constantes.

La solución para la ecuación diferencial 2.6 es:

$$B(t) = \frac{K}{1 + \left(\frac{K}{B_0} - 1\right) \exp(-r(t - t_0))} \quad (2.7)$$

En la Figura 2.2 se indican las trayectorias de la ecuación 2.7 según la relación entre la condición inicial (B_0) y K .

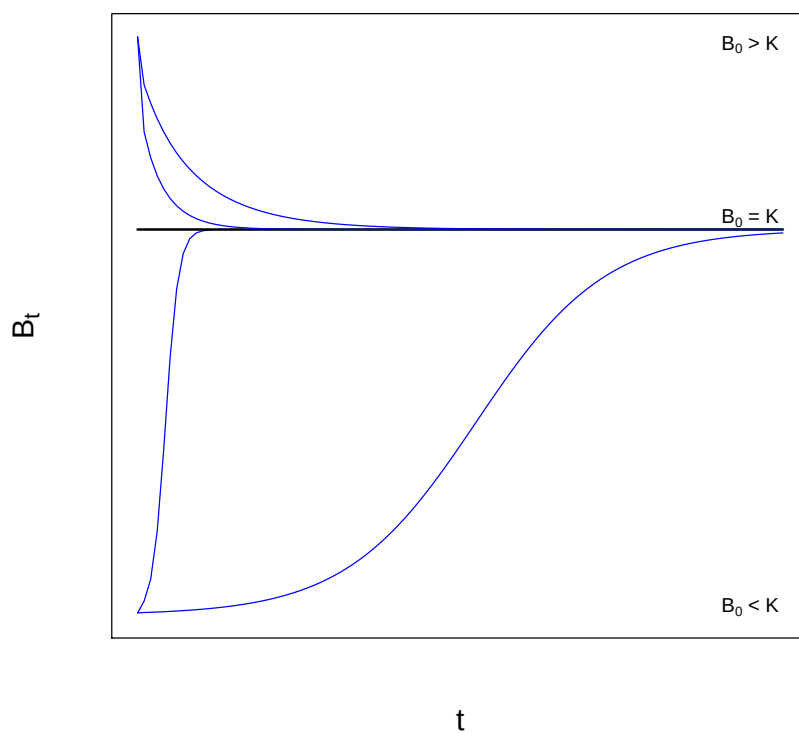


Figura 2.2: Trayectorias de las posibles soluciones de la ecuación logística 2.7 para el caso particular de r y K constantes y los casos críticos: $K < B_0$, $K = B_0$, $K > B_0$.

Ecuación logística con explotación proporcional y explotación constante

Los modelos más comúnmente aplicados son aquellos con tasa constante de explotación o con explotación denso-dependiente ([5] Pág. 305 a 310). Se incorpora a la ecuación logística el término de explotación denso-dependiente $H_p B$, donde H_p es la constante de proporcionalidad de explotación.

$$\begin{cases} \frac{dB(t)}{dt} = r(t)B(t)\left(1 - \frac{B(t)}{K(t)}\right) - H_p B(t) \\ B(t_0) = B_0 \end{cases} \quad (2.8)$$

Si $r(t) \neq 0$ entonces se puede reescribir 2.8 de la siguiente forma:

$$\begin{cases} \frac{dB(t)}{dt} = r'(t)B(t)\left(1 - \frac{B(t)}{K'(t)}\right) \\ B(t_0) = B_0 \end{cases} \quad (2.9)$$

donde $r'(t) = r(t) - H_p$ y $K'(t) = K(t)\frac{r'(t)}{r(t)}$. Para el caso particular de K' , r' y H_p constantes la solución a la ecuación 2.9 resulta:

$$B(t) = \frac{K'}{1 + \left(\frac{K'}{B_0} - 1\right) \exp(-r'(t - t_0))} \quad (2.10)$$

Las trayectorias de la solución 2.10 son similares a la de la ecuación 2.7 cuando $r' > 0$ y $B(t) \rightarrow K$ con $t \rightarrow +\infty$. Se observa que la capacidad de carga $K' = \left(1 - \frac{H_p}{r}\right)K$, por lo tanto la población se dirige a la extinción cuando $r' < 0$. Para más detalle ver [5], Sección 2.4 y 2.5).

Explotación constante

En este escenario se asume una tasa H_c de explotación constante, es decir que no depende de la Biomasa en el instante t . El sistema bajo este escenario es el siguiente.

$$\begin{cases} \frac{dB(t)}{dt} = r(t)B(t)\left(1 - \frac{B(t)}{K(t)}\right) - H_c \\ B(t_0) = B_0 \end{cases} \quad (2.11)$$

La solución de la ecuación diferencial 2.11, si $r(t)$ y $K(t)$ son constantes, se presenta a continuación.

$$B(t) = \frac{(D - r) + A(D + r) \exp(Dt)}{\frac{2r}{K}(A \exp(Dt) - 1)} \quad (2.12)$$

donde

$$\begin{aligned} D &= \sqrt{r^2 - \frac{4rH_c}{K}} \\ A &= \frac{r - D + J}{r + D + J} \exp(-Dt_0) \\ J &= \frac{-2rB_0}{K} \end{aligned} \quad (2.13)$$

Si $D > 0$ y r y K constantes, la ecuación 2.11 puede escribirse como.

$$\frac{dB(t)}{dt} = \frac{r}{K}(B(t) - B_1)(B(t) - B_2) \quad (2.14)$$

donde B_1 y B_2 son las raíces de 2.14

$$B_1 = \frac{K(r - D)}{2r}$$

$$B_2 = \frac{K(r + D)}{2r}$$

si r , K y H_c son positivos entonces $D < r$ y $0 < B_1 < B_2$ de manera que:

$$\frac{dB}{dt} > 0 \quad \text{para} \quad B_1 < B < B_2$$

$$\frac{dB}{dt} < 0 \quad \text{para} \quad B < B_1 \quad \text{o} \quad B > B_2$$

Se deduce que:

- si $B_0 > B_2$ entonces $B(t)$ decrece y tiene limite B_2 . Si $B_1 < B_0 < B_2$ entonces $B(t)$ crece y tiene limite B_2
- $B_0 < B_1$ entonces $B(t)$ decrece y se vuelve negativa y tiene asíntota vertical en el valor de t dado por $Dt = \log(\frac{1}{A})$ con t positivo

$$r - D + J > r - D - \frac{2rB_1}{K} = 0$$

$$r + D + J > r - D + J$$

El equilibrio positivo B_2 existe cuando $H_c < \frac{rk}{4}$ y por lo tanto $D > 0$. Pero si $B_0 < B_1$, la extinción ocurre porque la tasa de nacimiento es insuficiente para superar la explotación. Si $H_c > \frac{rk}{4}$ entonces D es imaginario y la forma cuadrática $(-\frac{r}{K})B^2 + rB - H_c$ tiene signo negativo para todo B , por esto $\frac{dB}{dt} < 0$ y $B(t)$ es negativo para cualquier valor de B_0 .

2.2.2. Ecuación en diferencias de crecimiento logístico con explotación

El supuesto de que el tamaño de las poblaciones varía en pequeños intervalos de tiempo es poco realista. Por lo que, para una gran población homogénea, donde los nacimientos y las muertes se producen continuamente a través del tiempo y las generaciones se superponen completamente, podría aproximarse la ecuación 2.8 con una ecuación en diferencias finitas. También es posible añadir los rezagos de tiempo explicitandolos en la ecuación diferencial. Sin embargo, para una población en un determinado ambiente en una determinada estación, especialmente con la no superposición de generaciones o cohortes, un modelo de tiempo discreto puede ser más apropiado ([8] Pág. 31). Por lo que

discretizando el modelo continuo se puede plantear el modelo de crecimiento logístico de la siguiente manera:

$$B_{t+1} = B_t + r_t B_t \left(1 - \frac{B_t}{K_t}\right) - C_t$$

2.3. Modelo de Schaefer

Este modelo desarrollado por Schaefer (1954) toma como base el modelo de crecimiento logístico de población desarrollado por Verhulst (1838). Usualmente en la literatura pesquera se conoce como un “modelo de producción excedentaria”. El modelo es el siguiente:

$$\begin{cases} B_{t+1} = B_t + r B_t \left(1 - \frac{B_t}{K}\right) - C_t \\ B_{t_0} = B_0 \end{cases} \quad (2.15)$$

En este modelo la tasa intrínseca de crecimiento r y la capacidad de carga K son constantes en el tiempo. Además Schaefer incorporó el supuesto de que la explotación, en este caso la captura (C) es una función lineal de la Biomasa y el esfuerzo de captura:

$$C_t = q E_t B_t \quad (2.16)$$

donde q es el coeficiente de capturabilidad y E_t es el esfuerzo de pesca, o tasa de explotación, en el instante t , y cuyas unidades se detallan en el capítulo 5.

El índice de abundancia, en adelante *Captura Por Unidad de Esfuerzo* CPUE, es asumido generalmente como proporcional a la Biomasa y esta dado por la siguiente ecuación:

$$CPUE_t = qB_t \quad (2.17)$$

A partir de estos parámetros biológicos se obtienen una serie de indicadores, denominados *puntos biológicos de referencia*, estos son:

- El rendimiento máximo sostenible RMS , bajo el supuesto de equilibrio que ($B_t = B_{t+1}$):

$$RMS = \frac{rK}{4} \quad (2.18)$$

- Tasa de rendimiento máximo sostenible RMS :

$$h_{RMS} = \frac{r}{2} \quad (2.19)$$

- Biomasa máxima sostenible o Biomasa $_{\acute{o}ptima}$:

$$B_{\acute{o}ptima} = \frac{K}{2} \quad (2.20)$$

- La Biomasa virgen es :

$$B_{virgen} = K \quad (2.21)$$

En la Figura 2.3 se observan los puntos anteriormente mencionados

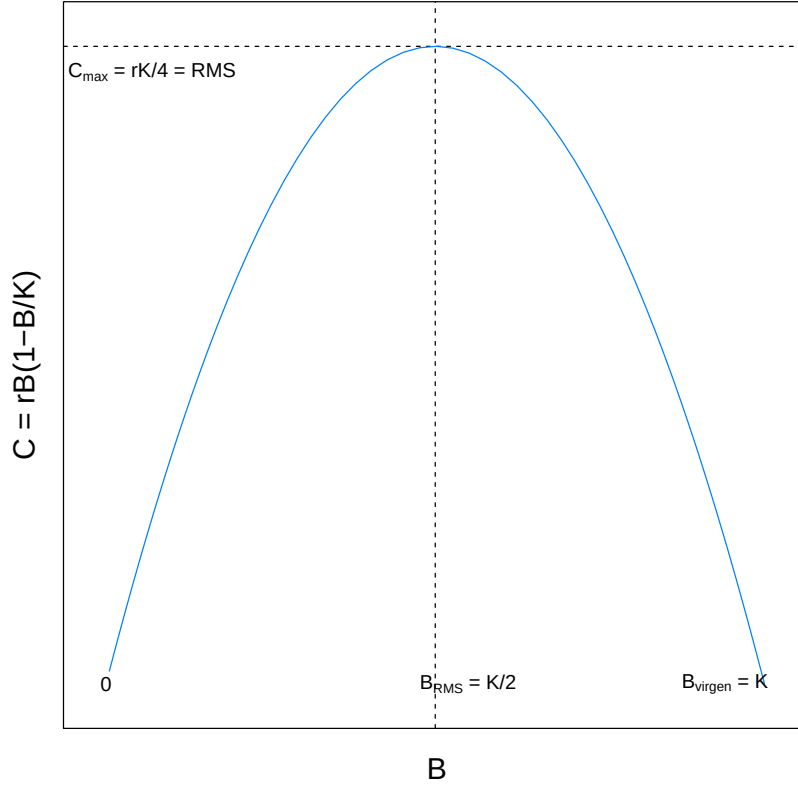


Figura 2.3: Modelo de Schaefer en equilibrio, $B_{t+1} = B_t$.

Incorporación de estocasticidad en el modelo

Hasta el momento el modelo de interés es determinístico. Para la incorporación de la incertidumbre en el mismo se tienen en cuenta las siguientes consideraciones:

- La aleatoriedad propia de los procesos de interés (tasa de nacimientos, tasa de muerte, etc), este tipo de incertidumbre se la conoce como *incertidumbre del proceso*, pues es la aleatoriedad de los parámetros (procesos) biológicos del modelo matemático.
- La incertidumbre propia de los errores de medición, denominada en la literatura como *incertidumbre en las observaciones*.

La incorporación de la aleatoriedad en el modelo determinístico de Schaefer, de acuerdo a Hilborn ([9], Cap. 7 y 10), se introduce de modo general como muestran las

ecuaciones 2.22 y 2.23.

$$B_{t+1} = \left[B_t + rB_t \left(1 - \frac{B_t}{K} \right) - C_t \right] \exp \left(W_t \sigma_w - \frac{\sigma_w^2}{2} \right) \quad (2.22)$$

$$CPUE_t = qB_t \exp \left(V_t \sigma_v - \frac{\sigma_v^2}{2} \right) \quad (2.23)$$

donde W_t y V_t son los términos que introducen la aleatoriedad en el proceso y en las observaciones respectivamente, se asume que tienen distribución normal con media 0 y desvío estándar 1. Como consecuencia los términos de errores del proceso $\exp \left(W_t \sigma_w - \frac{\sigma_w^2}{2} \right)$ y de las observaciones $\exp \left(V_t \sigma_v - \frac{\sigma_v^2}{2} \right)$ tienen distribución Log-Normal. La incorporación de incertidumbre a través de la forma multiplicativa evita problemas de valores negativos en los valores de la Biomasa y la $CPUE$.

Distribución Log-Normal

Características de la distribución Log-Normal:

- Si $X \sim Normal(\mu, \sigma^2) \rightarrow Y = e^{(X)} \sim \log - Normal(\mu, \sigma^2)$
- Función de densidad de probabilidad:

$$f_Y(y; \mu, \sigma) = \frac{1}{y\sigma\sqrt{2\pi}} e^{-(\ln(y)-\mu)^2/2\sigma^2} I_{\{y>=0\}}$$

- Función de distribución acumulada de probabilidad:

$$F_Y(y; \mu, \sigma) = \Phi \left(\frac{\ln y - \mu}{\sigma} \right)$$

donde Φ es la distribución normal estandarizada.

- Esperanza de Y

$$E(Y) = e^{\mu + \sigma^2/2}$$

- Mediana de Y

$$\text{Med}(Y) = e^\mu$$

- Varianza de Y :

$$\text{var}(Y) = (e^{\sigma^2} - 1)e^{2\mu + \sigma^2}$$

Pruebas de normalidad

Se desea probar las siguientes hipótesis

$$\begin{cases} H_0 : F(x) = F_0(x) & \text{contra} \\ H_1 : \text{no } H_0 \end{cases} \quad (2.24)$$

Donde $F_0(x)$ es la función de distribución normal.

Prueba de Shapiro - Wilk

Si un conjunto de observaciones, x_i , proviene de la distribución normal, entonces

$$x_i = \mu + \sigma z_i$$

Si $E(x_i) = w_i$ es la esperanza de i -ésimo estadístico de orden y $\mathbf{V}$ es la matriz de varianzas y covarianzas de los estadísticos de orden $x_1, x_2, \dots, x_n$, entonces el mejor estimador lineal insesgado de σ se obtiene de la regresión generalizada por mínimos cuadrados de la muestra de estadísticos de orden en sus valores esperados, que es

$$b = \mathbf{a}'\mathbf{x}$$

donde

$$\mathbf{a} = \frac{\mathbf{w}'\mathbf{V}^{-1}}{(\mathbf{w}'\mathbf{V}^{-1}\mathbf{V}^{-1}\mathbf{w})^{\frac{1}{2}}}$$

$\mathbf{a}'\mathbf{a} = 1$. En el vector $\mathbf{a}$, la componente $a_i = -a_{n-i+1}$ entonces b puede calcularse como

$$b = \sum_{i=1}^{[n/2]} a_{n-i+1}(x_{(n-i+1)} - x_{(i)}) \quad (2.25)$$

para valores adecuados de a_i . El estadístico de Shapiro - Wilk se define de la siguiente manera

$$W = \frac{b^2}{(n-1)s^2}$$

donde s^2 es la varianza muestral. Esta prueba funciona adecuadamente para muestras de tamaño menor a 50.

Prueba de Kolmogorov - Smirnov

Esta prueba compara la función de distribución empírica con la distribución hipotética. La función de distribución empírica de una muestra, $F_n(x)$, se define como.

$$F_n(x) = \begin{cases} 0 & x < x_{(1)} \\ \frac{i}{n} & x_{(i)} \leq x < x_{(i+1)} \quad i = 1, \dots, n-1 \\ 1 & x_{(n)} \leq x \end{cases} \quad (2.26)$$

y la función de distribución bajo la hipótesis nula, en este caso la normal, $p_{(i)} = \Phi\left(\frac{x_{(i)} - \hat{\mu}}{\hat{\sigma}}\right)$. El estadístico de esta prueba se basa en la máxima diferencia entre la función de distribución empírica, $F_n(x)$, y la función de distribución teórica de la hipótesis nula, $F_0(x)$, correspondiente a la distribución normal.

$$D = \max |F_n(x) - F_0(x)|$$

este estadístico se basa en la máxima diferencia entre la función de distribución empírica y la función de distribución normalidad, dadas por:

$$D^+ = \max_{i=1, \dots, n} \left[\frac{i}{n} - p_{(i)} \right]$$

$$D^- = \max_{i=1, \dots, n} [p_{(i)} - \frac{i-1}{n}]$$

$$D = \max[D^+, D^-]$$

2.4. Enfoque Bayesiano

2.4.1. Elementos de la Teoría Bayesiana

En este enfoque estadístico, se asume que los parámetros de interés son variables aleatorias con ciertas distribuciones de probabilidad ([7] Pág. 6 a 9). Si Θ es el vector de parámetros, según el Teorema de Bayes la distribución de probabilidad a posteriori de Θ es calculada de la siguiente forma:

$$p(\Theta|\mathbf{y}) = \frac{p(\Theta, \mathbf{y})}{p(\mathbf{y})} = \frac{p(\Theta)p(\mathbf{y}|\Theta)}{\int_{\Theta} p(\Theta)p(\mathbf{y}|\Theta)} \quad (2.27)$$

donde $p(\Theta)$ es la distribución a priori de los parámetros, $p(\mathbf{y}|\Theta)$ hace referencia a la distribución de los datos y $p(\Theta|\mathbf{y})$ representa la distribución a posteriori de Θ . Además $p(\Theta, \mathbf{y})$ representa la distribución conjunta de $\Theta, \mathbf{y}$.

También se puede plantear la distribución a posteriori no normalizada de Θ de la siguiente forma:

$$p(\Theta|\mathbf{y}) \propto p(\Theta)p(\mathbf{y}|\Theta) \quad (2.28)$$

En este trabajo la variable de interés y es la *CPUE* y el vector Θ esta formado por los parámetros r , K , q , el desvío de la variable *CPUE* dado por las observaciones (σ^2) y

el desvío de la variable $CPUE$ dado por el proceso (τ^2). Aquí hay que tener en cuenta un aspecto muy importante en este tipo de análisis: “La investigación sobre las distribuciones a priori es sin duda el aspecto más controvertido de cualquier análisis Bayesiano. Por lo tanto, recomendamos mucho cuidado al realizar una evaluación Bayesiana, dado que se debe tener en cuenta una investigación exhaustiva sobre varias distribuciones a priori. Este proceso debe incluir información que especifica qué modelos se consideraron para su inclusión en el análisis y por qué algunos de estos modelos no fueron considerados a pesar de que pueden ser plausibles” ([13] Pág. 42 a 43) . Teniendo en cuenta esto se pueden considerar distribuciones a priori usando criterios biológicos y en algunos casos asumir distribuciones no informativas para algunas variables.

Se incorpora incertidumbre al modelo determinista de Schaefer mediante variables aleatorias independientes e idénticamente distribuidas de forma multiplicativa que representan respectivamente un error de proceso y el error de observación. Se utiliza en este caso la reparametrización realizada por Meyer y Millar (ver más detalles en [12]) para realizar inferencia Bayesiana denominada **Modelo de espacios de estado** ([12] Pág. 1079 a 1080). Los espacios de estado se refieren a las observaciones de la serie de tiempo observada de $CPUE_t$ “estados” y a la serie de tiempo no observada B_t . Estos estados se suponen que siguen un modelo de transición estocástica.

Si se asume que la Biomasa anual es como un porcentaje de la capacidad de carga se puede plantear la siguiente reparametrización: $P_t = B_t/K$, el motivo de utilizar esta reparametrización es acelerar el muestreo de Gibbs ([12] Pág. 1080). Además se supone que la incertidumbre tiene una distribución Log-Normal con los parámetros que se detallan a continuación. Bajo estos supuestos el nuevo modelo se puede expresar de la siguiente forma:

$$\begin{cases} P_1 | \sigma^2 = \exp(u_1) \\ P_t | P_{t-1}, r, K, \sigma^2 = \left(P_{t-1} + r P_{t-1} (1 - P_{t-1}) - \frac{C_{t-1}}{K} \right) \exp(u_t) \\ CPUE_t | P_t, q, \tau^2 = (q K P_t) \exp(v_t) \end{cases} \quad (2.29)$$

donde u_t y v_t son variables aleatorias i.i.d. normales con media 0 y varianza σ^2 para

u_t y varianza τ^2 para v_t para cada instante t .

Se debe plantear todas las distribuciones a priori conjuntas para los cinco parámetros r , K , q , σ^2 y τ^2 y los estados desconocidos $P_1, \dots, P_N$, además la distribución conjunta de los índices de abundancia $CPUE_1, \dots, CPUE_N$. Se asume que los parámetros r , K , q , σ^2 y τ^2 son independientes a priori. Por una aplicación sucesiva del Teorema de Bayes y la independencia condicional de los estados siguientes, la densidad conjunta a priori queda dada por:

$$\begin{aligned} p(r, K, q, \sigma^2, \tau^2, P_1, \dots, P_N) &= \\ &= p(r)p(K)p(q)p(\sigma^2)p(\tau^2)p(P_1|\sigma^2) \times \prod_{t=2}^N p(P_t|P_{t-1}, r, K, \sigma^2) \end{aligned} \quad (2.30)$$

Las distribuciones a priori para r , K , q , σ^2 , τ^2 son definidas usando criterios biológicos. Por el supuesto de independencia condicional referida al índice de abundancia ($CPUE$) la distribución muestral o de los datos está dada por la siguiente expresión:

$$p(CPUE_1, \dots, CPUE_N | r, K, q, \sigma^2, \tau^2, P_1, \dots, P_N) = \prod_{t=1}^N p(CPUE_t | P_t, q, \tau^2) \quad (2.31)$$

Entonces, por el Teorema de Bayes, la distribución a posterior de los parámetros es proporcional a la distribución conjunta dada por la siguiente forma:

$$\begin{aligned} p(r, K, q, \sigma^2, \tau^2, P_1, \dots, P_N, CPUE_1, \dots, CPUE_N) &= \\ &= p(r)p(K)p(q)p(\sigma^2)p(\tau^2)p(P_1|\sigma^2) \times \prod_{t=2}^N p(P_t|P_{t-1}, r, K, \sigma^2) \times \prod_{t=1}^N p(CPUE_t | P_t, q, \tau^2) \end{aligned}$$

2.4.2. Muestreo de Gibbs

El Muestreo de Gibbs ([7] Pág. 287 a 305) consiste en generar muestras aleatorias a partir de las distribuciones condicionales completas para cada una de las variables que intervienen en el modelo. Disponemos de un conjunto de variables aleatorias $X_0, X_1, \dots, X_N$, y un conjunto de observaciones y . Dado un vector inicial $X_0^0, X_1^0, \dots, X_N^0$ obtenido mediante la simulación de las distribuciones a priori, el algoritmo sigue los siguientes pasos:

Paso 1:

1^{era} realización: $X_0^{(1)} \sim f(X_0|X_1^{(0)}, \dots, X_N^{(0)}, y)$

2^{da} realización: $X_1^{(1)} \sim f(X_1|X_0^{(1)}, X_2^{(0)}, \dots, X_N^{(0)}, y)$

...

realización final : $X_N^{(1)} \sim f(X_N|X_0^{(1)}, X_1^{(1)}, \dots, X_{N-1}^{(1)}, y)$

Paso 2:

Tomando como vector inicial: $X_0^{(1)}, X_1^{(1)}, \dots, X_N^{(1)}$ se vuelve a realizar el paso 1 nuevamente. Se puede demostrar que esta secuencia de valores generados constituyen una cadena de Markov (ver capítulo 11 de [7]), de forma que operando de forma sucesiva un número suficientemente grande de veces (n), se obtendrá el vector $X_0^{(n)}, X_1^{(n)}, \dots, X_N^{(n)}$ que converge en distribución a una muestra aleatoria de la distribución conjunta $f(X_0, X_1, \dots, X_N)$ a medida que n tiende a infinito. Repitiendo el proceso G veces se obtendrá una muestra de tamaño G de extracciones aleatorias de la distribución conjunta, $X_0^{(j)}, X_1^{(j)}, \dots, X_N^{(j)}$ donde $j = 1, \dots, G$ que puede ser utilizada para estimar las densidades marginales y cualquier estadístico de dichas distribuciones.

2.4.3. Método MCMC y BUGS

Se realiza una representación utilizando diagramas DAG (Directed Acyclic Graph) ([12] Pág. 1080 a 1081) para indicar como se procede con la estimación a través del enfoque Bayesiano. Este diagrama (figura 2.4) tiene la ventaja de que permite concentrarse en la estructura esencial del modelo sin meterse en todos los detalles sobre las densidades. Para cualquier instante t , se representan todas las características no observables, los

parámetros desconocidos (r , K , q , σ_{obs} , σ_{proc} , P_t) y observables ($CPUE_{obs}$) como elipses y la única constante (Captura) en forma de rectángulo. Una forma de expresar el supuesto de independencia condicional es por un dibujo de las flechas sólidas entre los nodos, por ejemplo de $itau2$ a $CPUE[t]$. Las flechas dobles van a los nodos deterministas, que son las funciones lógicas de los demás nodos. El logaritmo de la media condicional de P_t ($Pmed[t]$), es un ejemplo de un nodo determinista, ya que es una función de los nodos K , r , la $Captura_{t-1}$, y P_{t-1} . En el diagrama la variabilidad en el proceso se tiene en cuenta a través de $isigma2$ e $itau2$ hace referencia a la variabilidad en las observaciones.

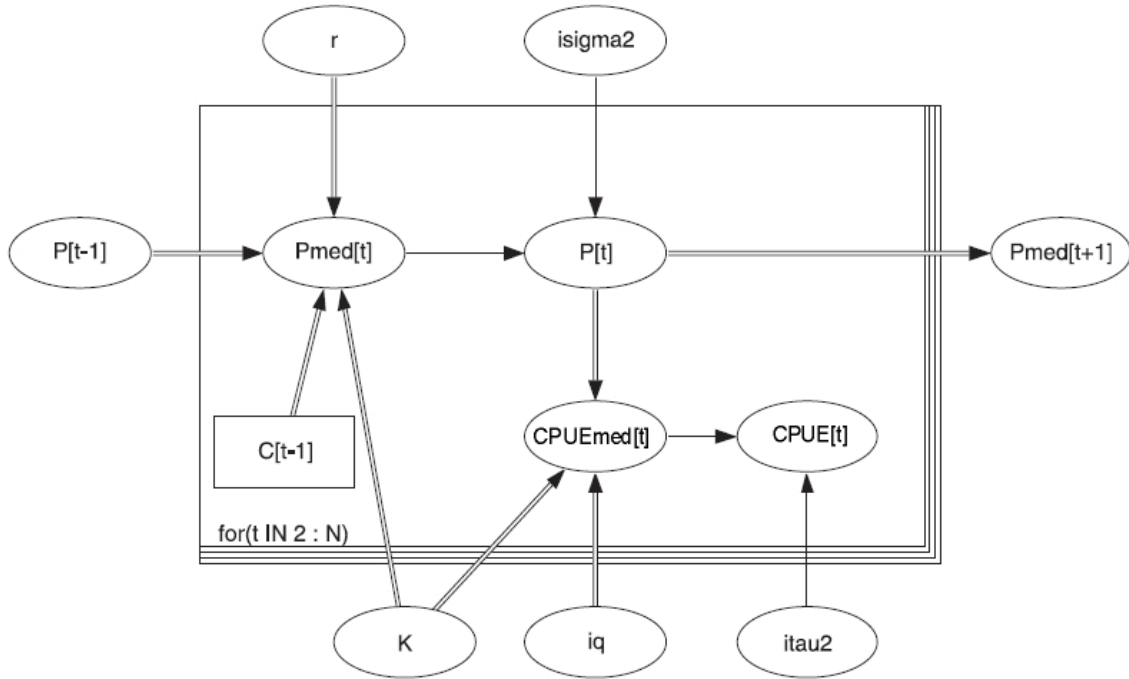


Figura 2.4: Representación del modelo de producción como un Directed Acyclic Graph (DAG).

La Figura 2.4 es la representación del modelo como DAG, por el supuesto de independencia condicional todos los grafos son dirigidos y no hay ciclos. V denota el conjunto de todos los nodos en el gráfico. Si $v \in V$ entonces v son llamados "padres" y predecesores son los llamados "hijos". Las flechas continuas indican que, dado el nodo padre, cada nodo v es independiente de todos los otros nodos excepto descendientes de v . Por

ejemplo, si se está en el instante t y se sabe la Biomasa en el instante $t - 1$ y los valores de los parámetros r , K , y *sigma.proc*, entonces la creencia en P_t es independiente de la Biomasa en instantes anteriores del 1 al $t - 2$ y los datos de todos los demás instantes, excepto el índice de abundancia relativa actual *CPUE* en el tiempo t .

Es entonces es fácil construir la distribución de probabilidad conjunta $p(v)$ de todos los nodos estocásticos con la descripción gráfica de los supuestos de independencia condicional:

$$p(v) = \prod_{v \in V} p(v | \text{padres}(v)) \quad (2.32)$$

Si V_u es el subconjunto de nodos no observables y el V_o es el subconjunto de nodos observables. Una vez que $p(V)$ se ha obtenido de la ecuación 2.32, una dificultad técnica que se encuentra en cualquier aplicación de la inferencia Bayesiana es obtener la constante de normalización para definir la distribución a posteriori de la variables no observables, para este caso viene dada por la ecuación:

$$p(V_u | V_o) \frac{p(V_u, V_o)}{p(V_o)} = \frac{p(V)}{\int_{V_u} p(V_u, V_o) dV_u} \quad (2.33)$$

Para el modelo planteado requiere $(N + 5)$ dimensiones de integración, ya que hay que integrar a los no observables: r , K , q , σ^2 , τ^2 , $P_1, \dots, P_N$. El cálculo de la distribución marginal a posteriori de cualquier variable requiere de una integración con $(N + 4)$ dimensiones. Esta cantidad elevada de dimensiones par integrar se puede resolver a través de Cadenas de Markov Monte Carlo (MCMC). El Muestreo de Gibbs es un algoritmo MCMC especial, que genera una muestra de la ecuación 2.33 e iterativamente toma muestras de cada una de las distribuciones condicionales completos a posteriori. Estas distribuciones condicionales completas a posteriori $p(v | V/v)$, para $v \in V_u$, pueden ser construidas a partir de la distribución condicional a posteriori $p(V)$ en la ecuación 2.32 con los términos que dependen de v :

$$p(v|V/v) \propto p(v|padres(v)) \times \prod_{w \in padres \in (v)} p(w|padres(w)) \quad (2.34)$$

Esto es facilitado por la representación DAG, como la distribución condicional completa a posteriori de cualquier nodo v sólo depende de sus padres, hijos y parientes. Por ejemplo, si $v = P_t$, entonces la distribución condicional completa a posteriori de P_t es: $p(P_t|r, K, q, \sigma^2, \tau^2, P_1, \dots, P_N, CPUE_1, \dots, CPUE_N)$, es proporcional a:

$$p(P_t|P_{t-1}, r, K, \sigma^2) \times p(P_{t-1}|P_t, r, K, \sigma^2) \times p(CPUE_t|P_t, q, \tau^2) \quad (2.35)$$

En este caso, hay dependencia de los nodos deterministas $Pmed_t$ y $CPUEmed_t$ dados por las funciones lógicas P_t , r , K y q . De esta manera, BUGS explora la representación del modelo como el DAG de la figura para la construcción de las distribuciones condicionales completas a posteriori de todos los nodos no observables. Una vez logrado esto, utiliza ciertos métodos de muestreo para tomar muestras de estas densidades. Para mas detalle del funcionamiento de BUGS ver [12] y [17].

Capítulo 3

Antecedentes

En una búsqueda exhaustiva en la bibliografía se pueden encontrar innumerables referencias sobre Modelos Dinámicos aplicados a recursos pesqueros. De los cuales referidos específicamente al tema del presente informe se pueden destacar los siguientes:

- “Confrontación de modelos de la Biomasa Dinámica: una aplicación a la Pesquería del Huachinango” [6]. En el mismo se describe con detalle los procedimientos para realizar la implementación de los seudocódigos planteados en el Capítulo 10 del libro “The Ecological Detective: Confronting Models with Data” [9].
- “Risk analysis of a Flatfish stock complex” [11], donde se puede encontrar un análisis similar al realizado en este trabajo, utilizando entre otros el modelo de Schaefer ajustado por la metodología de máxima verosimilitud y posteriormente haciendo un análisis Bayesiano del mismo.

Por otro lado no se encontró a nivel nacional un material referido específicamente al trabajo realizado en el presente informe.

Capítulo 4

Metodología

4.1. Base de datos de la literatura

Para ajustar el modelo se utiliza en primera instancia los datos referidos a la especie *Merluccius capensis*, merluza de Namibia ([9] Pág. 240).

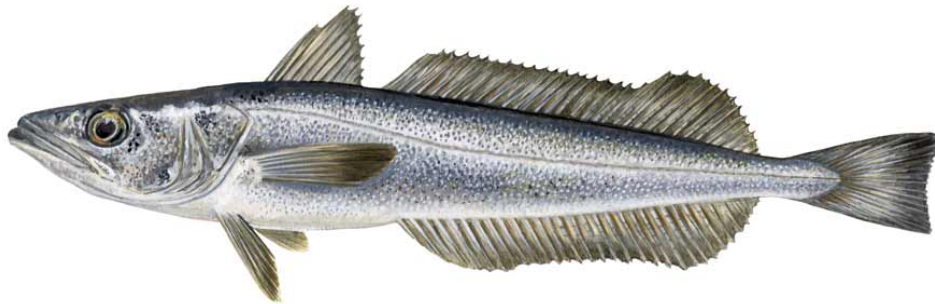


Figura 4.1: *Merluccius capensis*, merluza de Namibia.

El conjunto de datos utilizados se observa en la Figura 4.2.

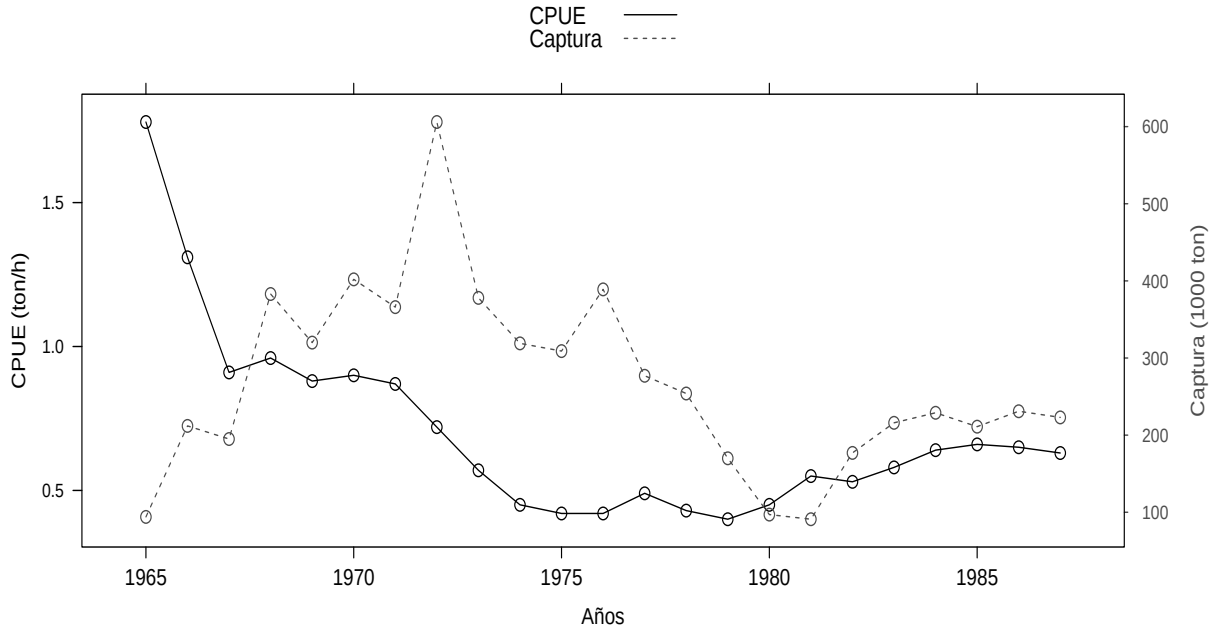


Figura 4.2: Evolución de la *CPUE* (ton/h) y *Captura* (10^3 ton) observada de *Merluccius capencis* en el periodo 1965-1987.

4.2. Base de datos de la Pesquería Uruguaya

Los datos utilizados en esta investigación provienen de la Dirección Nacional de Recursos Acuáticos - Ministerio de Ganadería Agricultura y Pesca y corresponden a la especie *Cynoscion guatucupa*, pescadilla.

La pescadilla (Figura 4.3) es un pez demersal de la clase Osteichthyes, orden Perciformes, familia Scianidae, que ocupa el tercer lugar en los desembarques de la República Oriental del Uruguay después de la merluza (*Merluccius hubbsi*) y la corvina (*Micropogonias furnieri*) y, precedido por la corvina, resulta el segundo en orden de importancia respecto a la pesca al arrastre costera [1].



Figura 4.3: *Cynoscion guatucupa*, pescadilla.

El conjunto de datos utilizados se observa en la Figura 4.4.

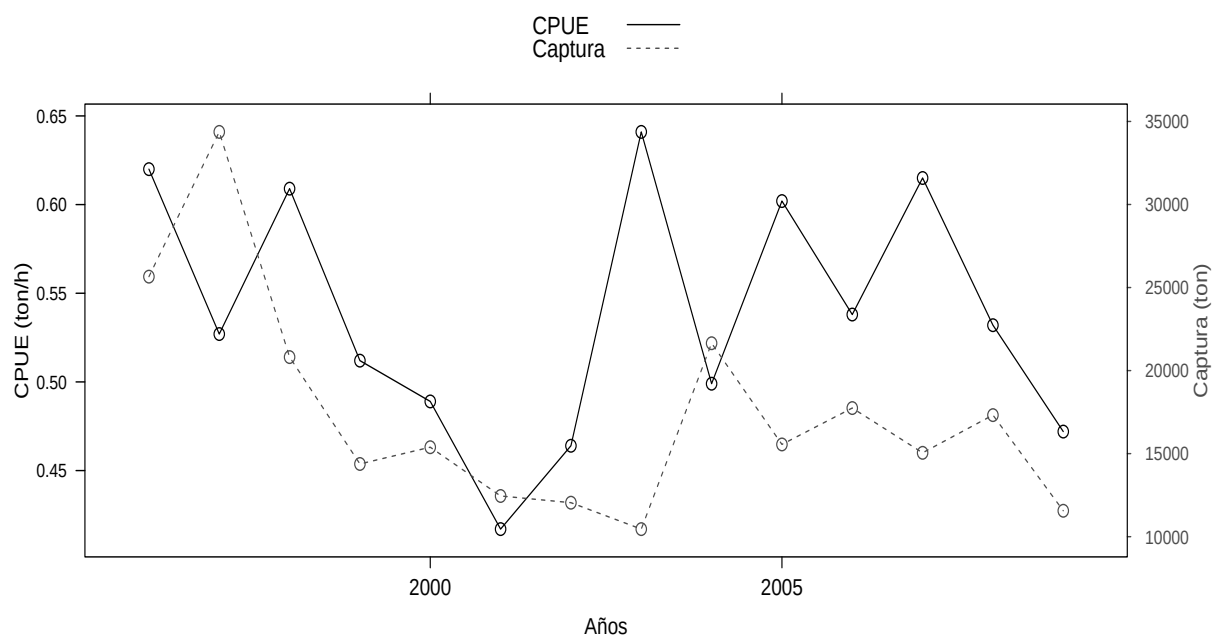


Figura 4.4: Evolución de la *CPUE* (ton/h) y *Captura* (ton) observada de *Cynoscion guatucupa* en el periodo 1996-2009.

Otras consideraciones

- Una forma usual sugerida por Hilborn ([9] Cap. 10) para el valor de la Biomasa inicial B_0 es tomarla igual a la capacidad de carga K . Pero no siempre se puede realizar este supuesto.

- El ajuste del modelo consiste en encontrar estimaciones para los parámetros r y K de forma que minimicen el logaritmo de la verosimilitud del modelo.

4.3. Descripción del procedimiento para ajustar el modelo usando Máxima Verosimilitud

El siguiente procedimiento está basado en los pseudocódigos planteados en Capítulo 10 del libro “The Ecological Detective: Confronting Models with Data” [9].

1. Se cargan los datos referidos a las series de tiempo de *CPUE* y *Captura*.
2. Ingresar los valores iniciales para del modelo: r , K , B_0 . Hay que tener la precaución de tomar valores iniciales de los parámetros de forma tal que el modelo pueda funcionar numéricamente. El valor de B_0 es clave para que el procedimiento pueda operar numéricamente y para su determinación se tiene en cuenta la información a priori sobre la pesquería en estudio, pero es difícil determinar su valor.
3. Para hacer un análisis exploratorio se gráfica la función de verosimilitud a través de superficies de nivel, mediante la cual se visualizan los pares de valores de r y K que podrían optimizar el modelo.
4. Se ejecuta la optimización de la suma en el tiempo t de la función de verosimilitud L_t como se indican en los esquemas 4.5 y 4.6, por lo que se busca encontrar los valores de los parámetros r y K que minimizan la suma de la función L_t .

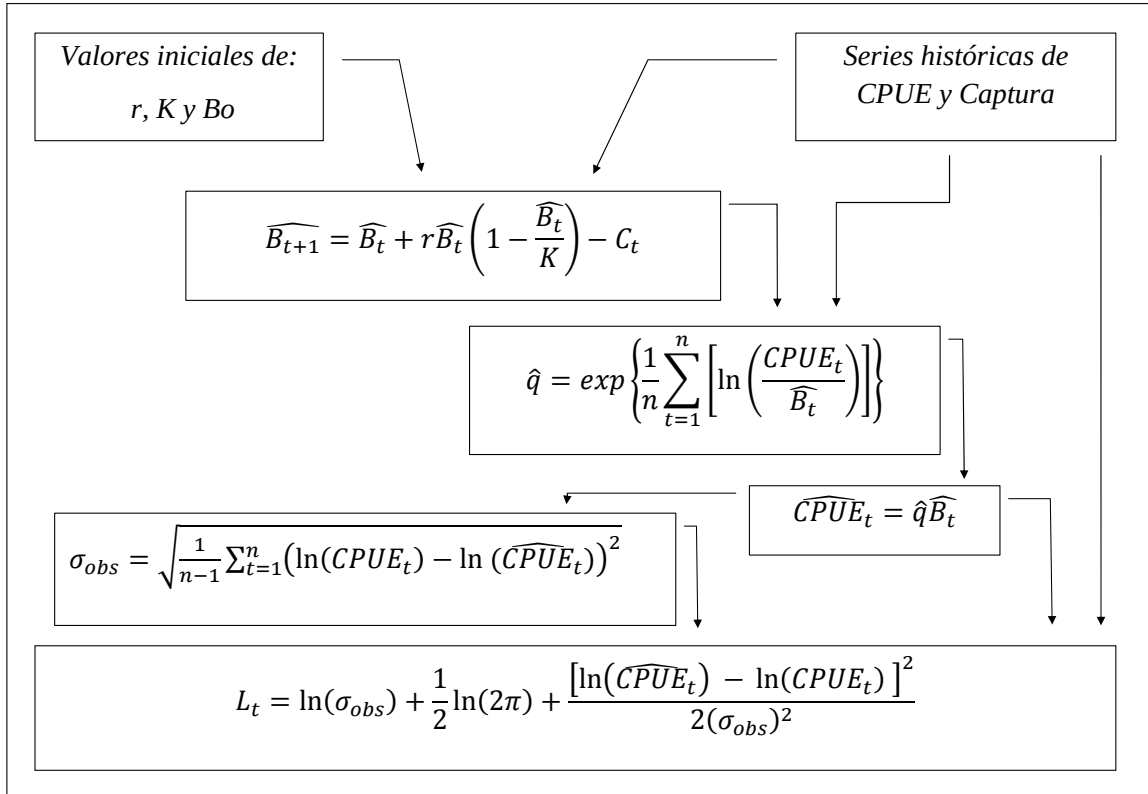


Figura 4.5: Esquema del procedimiento con error en las observaciones.

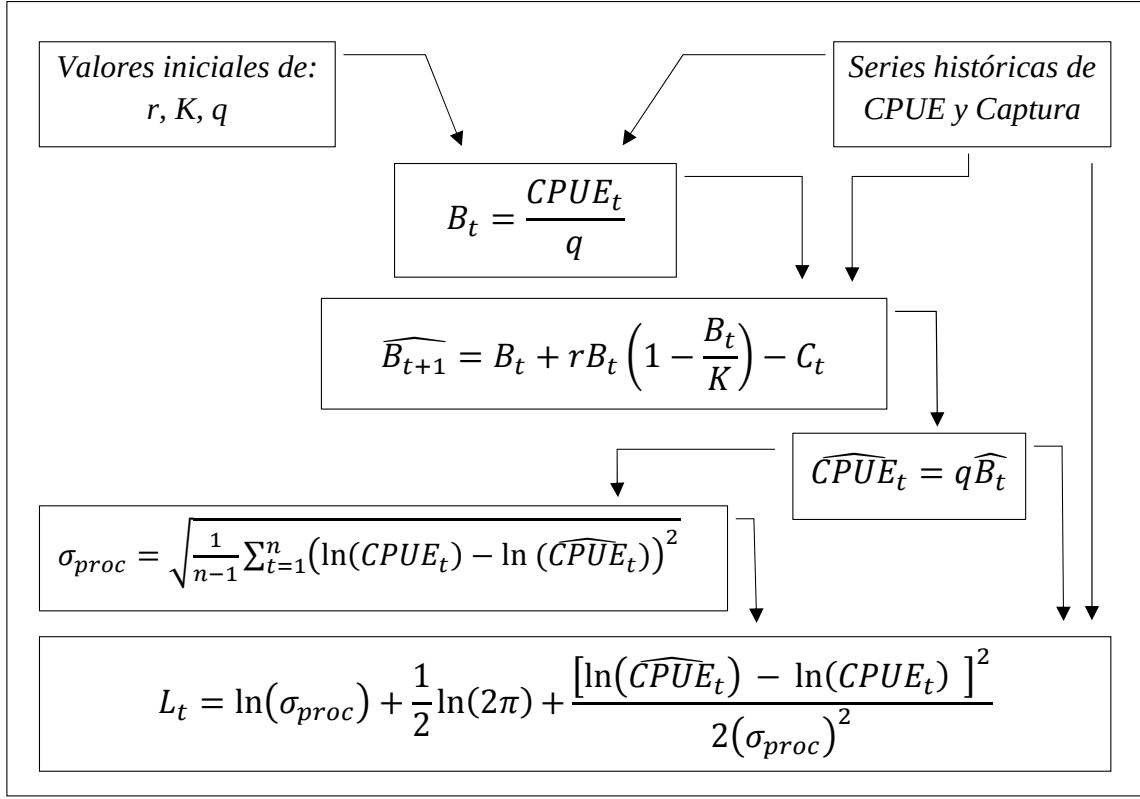


Figura 4.6: Esquema del procedimiento con error en el proceso.

5. Bajo el supuesto de incertidumbre en las observaciones se obtienen los valores óptimos de r y K , con estos se calcula: la serie de Biomasa predicha por el modelo, el coeficiente de capturabilidad y el valor de RMS .
6. Bajo el supuesto de incertidumbre en el proceso se obtienen los valores óptimos de r , K y q , con estos se calcula: la serie de Biomasa predicha por el modelo y el valor de RMS .
7. Se observa el ajuste del modelo con las gráficas de $CPUE$ y Biomasa en función del tiempo.
8. Se realizan las pruebas de normalidad (Shapiro - Wilk y Kolmogórov - Smirnov) para la diferencia de las variables logaritmo de la $CPUE$ y logaritmo de la $CPUE$ predicha por el modelo.

9. Una vez conocida la zona de optimización se construyen perfiles de verosimilitud para los parámetros r , K y el punto de referencia biológico RMS . El perfil de verosimilitud es el valor que toma la suma de la función logaritmo de la verosimilitud en el rango de valores de un parámetro donde se minimiza dicha función respecto a los parámetros restantes. Por ejemplo el perfil de verosimilitud para el parámetro r se obtiene de la siguiente forma:
 - a) Se define un rango de valores para el parámetro r .
 - b) Dado un valor inicial de K , se optimiza la suma de la función logaritmo de la verosimilitud para los distintos valores de r . Por lo que se obtiene para cada valor del rango de r , el valor de K que minimiza la función de interés y el valor que toma dicha función.
 - c) Con esta información se construyen intervalos de confianza para r utilizando el criterio de contrastes de verosimilitudes que establece que dos veces la diferencia entre la verosimilitud que mejor optimiza y otra verosimilitud tiene aproximadamente distribución Chi-cuadrado con un grado de libertad.
10. Para el caso del perfil de verosimilitud del punto de referencia RMS se debe tener en cuenta que este surge del cálculo $RMS = \frac{rK}{4}$ por lo tanto se considera un rango de valores para el RMS y un valor inicial para r (o K) y se busca el valor óptimo de r que hace mínima la verosimilitud perfil de RMS para cada uno de los valores del rango considerado, no es necesario dar un valor inicial para K ya que este se calcula como $\frac{4RMS}{r}$.
11. El perfil de verosimilitud y el intervalo de confianza para la capacidad de carga, K , se realizan de manera análoga al de r .

4.4. Programación en R

Para implementar procedimiento planteado anteriormente se utilizó el programa R Versión 2.15.0. [16]. Además se utilizaron las bibliotecas: Lattice y LatticeExtra. Los

códigos en R utilizados para realizar el análisis se incluyen en los anexos.

4.5. Programación en OpenBUGS

Para ajustar el modelo utilizando el enfoque Bayesiano se utilizó el programa de código abierto OpenBUGS. Los paquetes BUGS se usan para realizar Inferencia Bayesiana. Se especifica un modelo estadístico de cualquier complejidad, mediante la simple declaración de las relaciones entre las variables relacionadas. El programa incluye un "sistema experto", que determina una programación adecuada MCMC (Markov Chain Monte Carlo) basado en el muestreo de Gibbs, para analizar el modelo especificado [17]. Los códigos de OpenBUGS utilizados para realizar el análisis se incluyen en los anexos.

Capítulo 5

Resultados

5.1. Aplicación con datos de la literatura

5.1.1. Unidades

Se presenta en primera instancia el ajuste del modelo utilizando los datos referidos a la especie *Merluccius capensis*, merluza de Namibia ([9] Pág. 240).

El parámetro K y la Biomasa se miden en miles de toneladas, la tasa intrínseca de crecimiento r no tiene unidad, el coeficiente de capturabilidad q se mide en una proporción de 1/1000h y el Rendimiento Máximo Sostenible RMS se mide en miles de toneladas.

5.1.2. Ajuste del modelo de Schaefer por Máxima Verosimilitud con incertidumbre en las observaciones

En este apartado se presentan los resultados obtenidos del ajuste del modelo dinámico de Schaefer bajo el supuesto de incertidumbre en las observaciones.

Función de verosimilitud

Para tener una idea de la función objetivo que se optimiza, se gráfica la función del logaritmo de la verosimilitud normalizada (Figura 5.1 y 5.2). Para ello se toma un rango de valores para los parámetros r y K entre 0.2 y 0.6, y entre 2000 y 4000, respectivamente. La Figura 5.1 representa en tres dimensiones los valores que toma la función de interés para las distintas combinaciones de r y K .

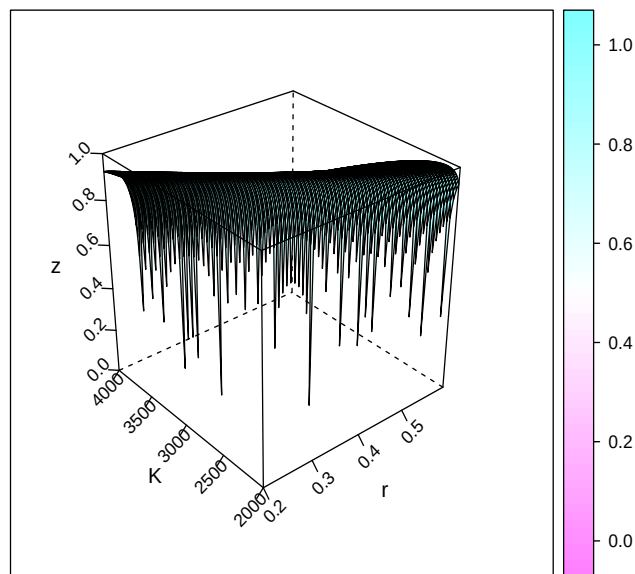


Figura 5.1: Logaritmo de la verosimilitud normalizada versus r y K .

La Figura 5.2 permite visualizar la región donde se optimiza (minimiza) la función de interés logaritmo de la verosimilitud normalizada.

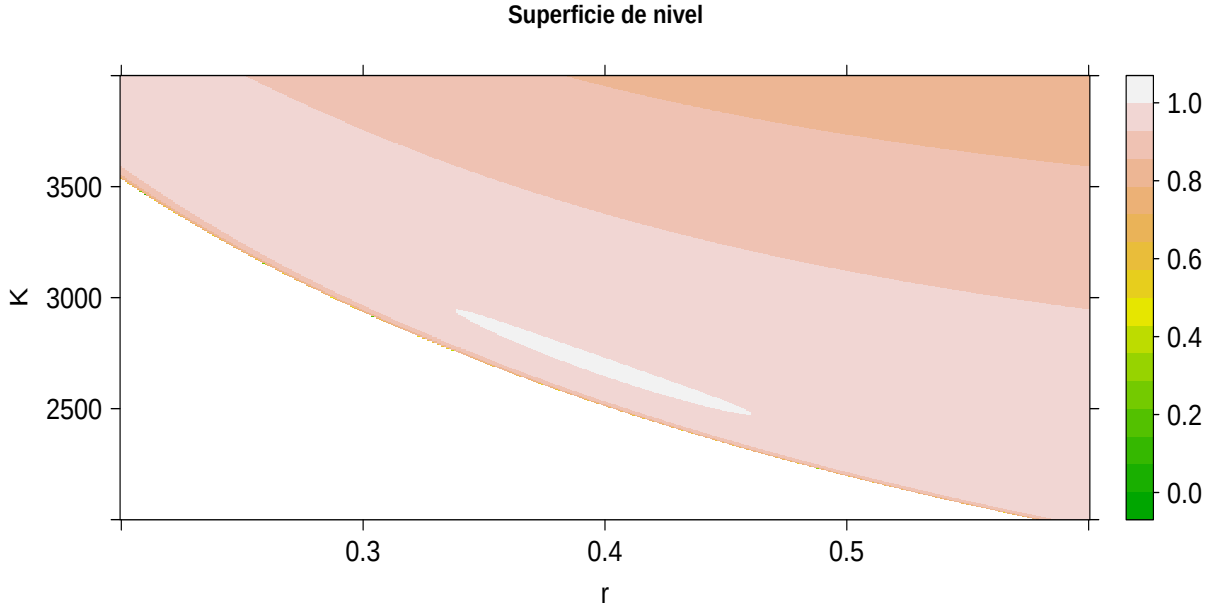


Figura 5.2: Superficie de nivel de la función de verosimilitud según los parámetros r y K .

La zona mas clara de la Figura 5.2, donde la función objetivo toma el valor uno (por estar normalizada), es donde se encuentra la región de valores óptimos de r y K .

Para encontrar los valores puntuales de r y K que minimizan la función del logaritmo de la verosimilitud de los datos, se tomaron valores iniciales de r entre 0.1 y 1 y valores iniciales de K entre 1000 y 10000, el resultado de aplicar el algoritmo de optimización es similar para todas las combinaciones posibles de estos dos parámetros,. La Tabla 5.1 contiene el resumen descriptivo del resultado del mecanismo de optimización para el modelo de Schaefer.

Parámetro/Resumen	Mín.	1 ^{er} Cuart.	Mediana	Media	3 ^{er} Cuart.	Máx.
r inicial	0.1000	0.3250	0.6000	0.5805	0.8000	1.0000
K inicial	2000	4000	6000	6341	8000	10000
r óptimo	0.393	0.394	0.394	0.394	0.394	0.394
K óptimo	2708	2709	2709	2709	2709	2711
$-sumLL$	-15.56	-15.56	-15.56	-15.56	-15.56	-15.56
RMS óptimo	266	267	267	267	267	267

Tabla 5.1: Estadística descriptiva del ajuste del modelo

Se obtuvo un rango acotado de valores óptimos para los parámetros r y K a pesar

de la variabilidad considerada para los valores iniciales de ambos. En la Tabla 5.2 se presentan algunos de los valores iniciales considerados para el ajuste del modelo y los valores óptimos obtenidos.

Valores iniciales		Óptimos						
r	K	r	K	σ_{obs}	q	RMS	$B_{\acute{o}ptima}$	$sum(-LL)$
0.4	8000	0.394	2709	0.125	0.00045	267	1355	-15.55
0.8	5000	0.394	2708	0.125	0.00045	267	1354	-15.55
1	10000	0.394	2709	0.125	0.00045	267	1355	-15.55

Tabla 5.2: Valores iniciales y valores óptimos

Con los valores óptimos de $r = 0.394$ y $K = 2709$, se obtiene el ajuste del modelo de Schaefer de Biomasa y CPUE, dicho ajuste se observa en la figura 5.3 y 5.4.

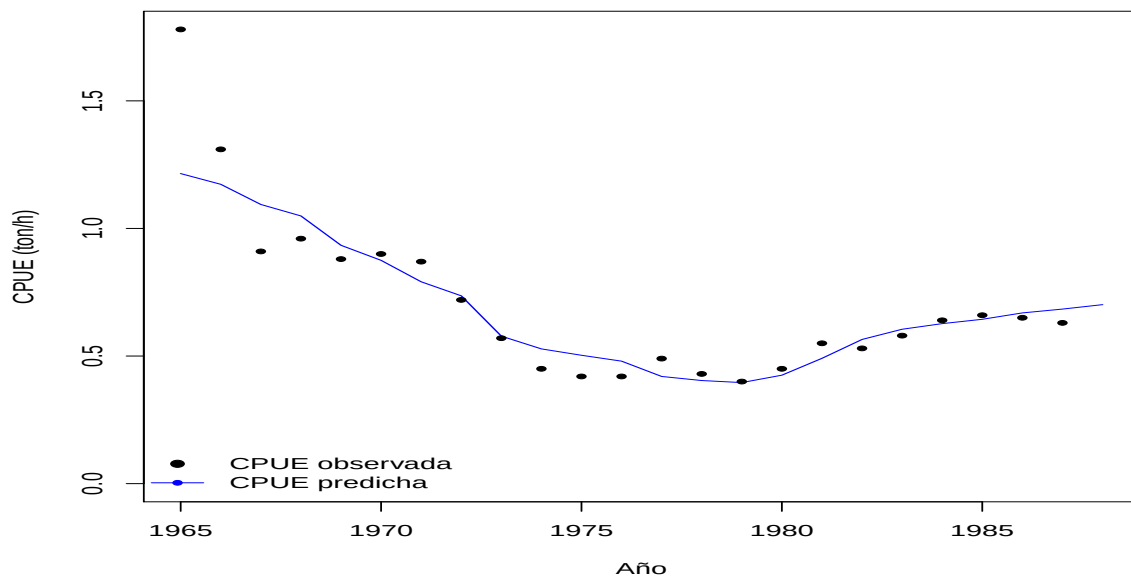


Figura 5.3: Evolución de la $CPUE$ observada y la $CPUE$ predicha por el modelo.

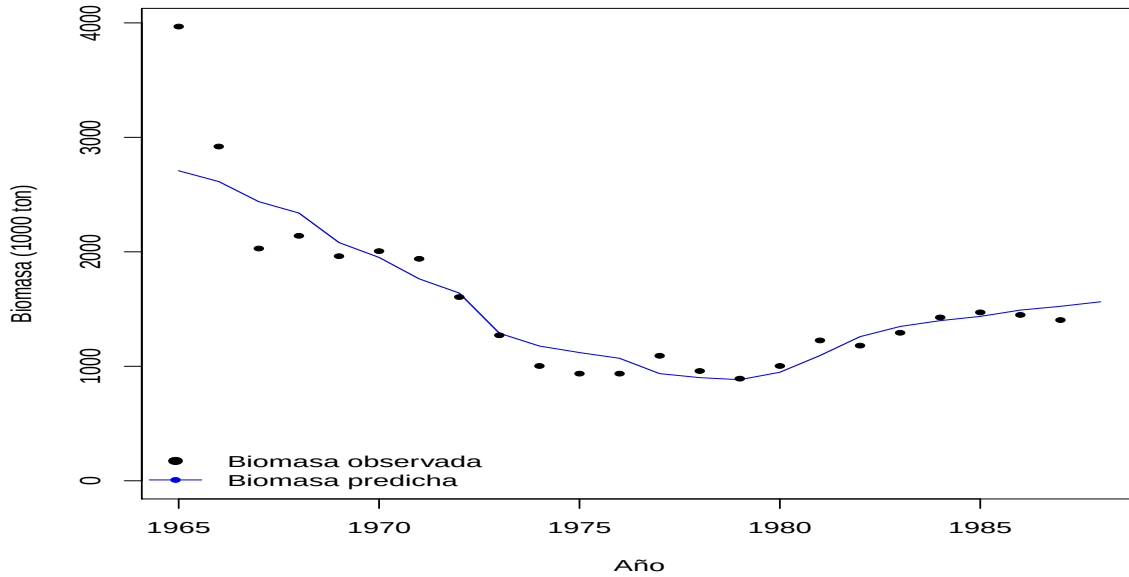


Figura 5.4: Evolución de la Biomasa observada y la Biomasa predicha por el modelo.

Resultados de las pruebas de normalidad

Para evaluar la normalidad del residuo estimado del modelo, se aplicaron las pruebas de normalidad de Shapiro - Wilk y Kolmogoróv - Smirnov.

En la tabla 5.3 se presenta el p-valor de las pruebas de normalidad al 5% de significación para la diferencia entre el logaritmo de la $CPUE$ observada y el logaritmo de la $CPUE$ ajustada.

Residuo estimado/Prueba	Shapiro - Wilk	Kolmogoróv - Smirnov
$\ln(CPUE_{Obs}) - \ln(CPUE_{Pred})$	0.25	0.97

Tabla 5.3: Pruebas de Normalidad

En ambas pruebas el p-valor es mayor que el nivel de significación por lo tanto no existe evidencia para rechazar la hipótesis de normalidad de los residuos del modelo.

Perfiles de verosimilitud

El perfil de la función de verosimilitud, permite identificar un rango de valores, para cada parámetro del modelo r y K y para el punto de referencia biológico RMS , donde se optimiza la función de verosimilitud, independientemente del valor que toman los parámetros restantes. Para ello se considera un rango de valores para cada parámetro y se optimiza la función logaritmo de la verosimilitud perfil del mismo respecto al restante. Por ejemplo para un rango de valores de r se evalúa que valor de K minimiza el logaritmo de la verosimilitud perfil del parámetro r . Las Figuras 5.5, 5.6 y 5.7 representan el perfil de verosimilitud para cada uno de los parámetros anteriormente mencionados.

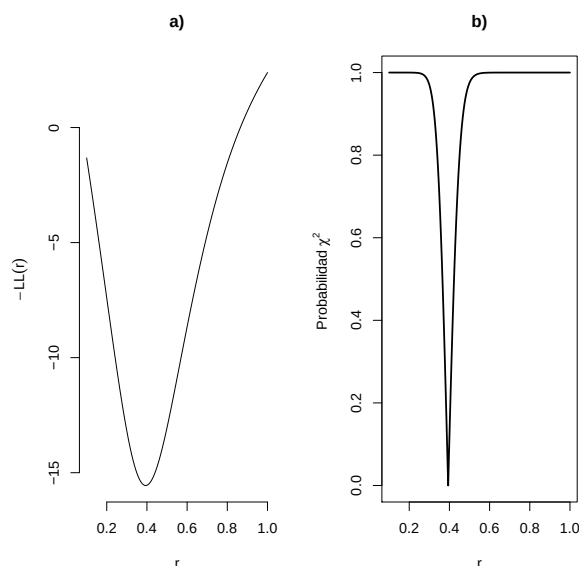


Figura 5.5: a) Perfil de verosimilitud de r ; b) Probabilidad Chi-cuadrado asociada a r .

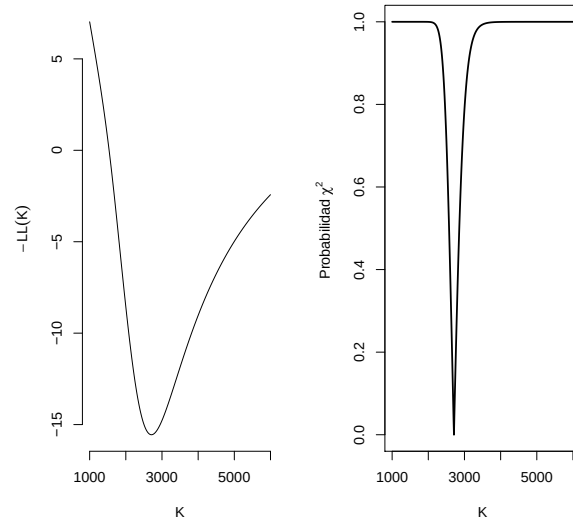


Figura 5.6: a) Perfil de verosimilitud de K ; b) Probabilidad Chi-cuadrado asociada a K .

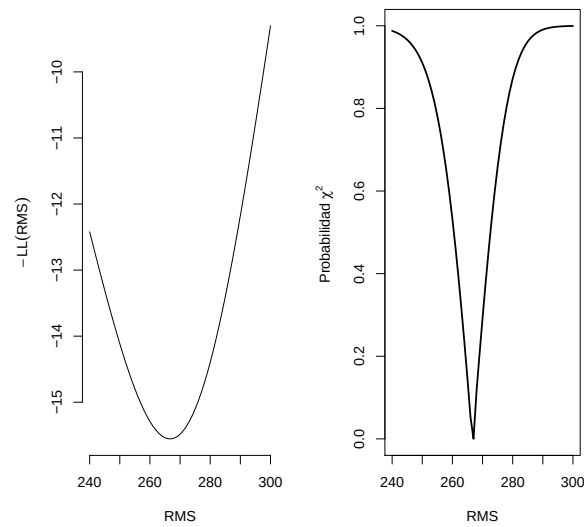


Figura 5.7: a) Perfil de verosimilitud del RMS ; b) Probabilidad Chi-cuadrado asociada al RMS .

Mediante las Figuras 5.5, 5.6, 5.7 y utilizando el criterio de la razón de verosimilitud planteado en el procedimiento, se obtienen los intervalos de confianza al 95 % para los

parámetros de interés:

Parámetro	Límite inferior	Límite superior
r	0.31	0.48
K	2390	3210
RMS	247	283

5.1.3. Ajuste del modelo de Schaefer por Máxima Verosimilitud con incertidumbre en el proceso

Al asumir incertidumbre en el proceso, la estimación de la Biomasa, mediante el modelo, depende de la $CPUE$ observada más que de la Biomasa estimada por el modelo en el año anterior, de modo que no se necesita estimar un valor inicial de Biomasa.

Al igual que bajo el supuesto de incertidumbre en las observaciones las Figuras 5.8, 5.9, 5.10 representan la superficie de nivel que permite visualizar la región donde la función objetivo se optimiza, esta indica el valor que toma el logaritmo de la función de verosimilitud para distintos pares de valores de los parámetros r y K , y para distintos q . En las gráficas siguientes (5.8, 5.9, 5.10) la zona más clara muestra los pares de valores de r y K que minimizan la función objetivo (normalizada).

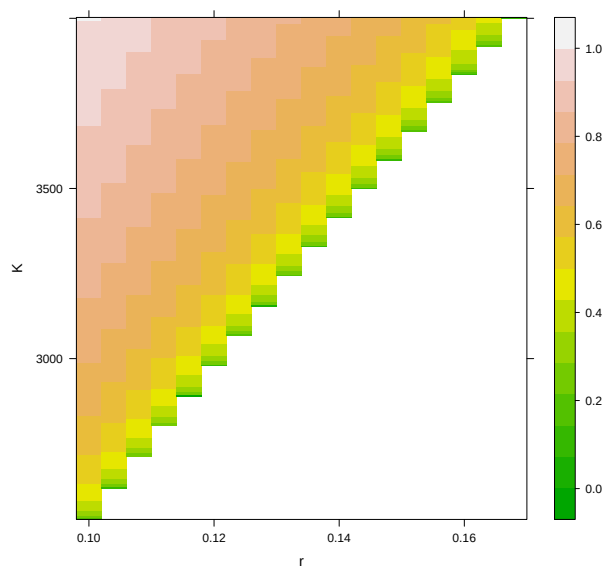


Figura 5.8: Curvas de nivel del logaritmo de la función de verosimilitud versus r y K para $q=6.42e-5$.

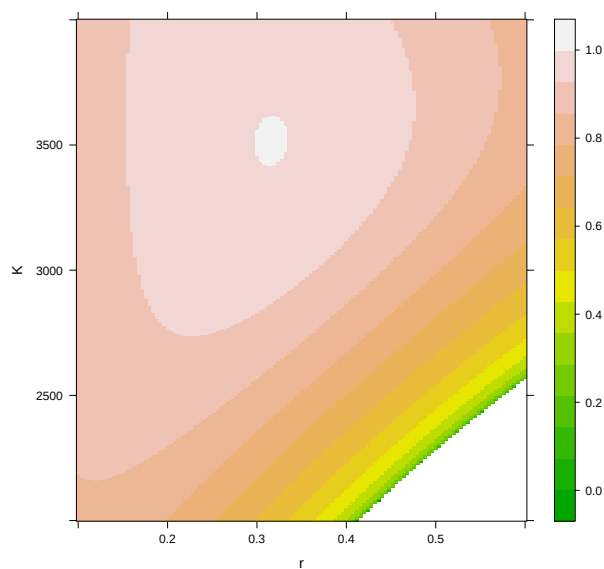


Figura 5.9: Curvas de nivel del logaritmo de la función de verosimilitud versus r y K para $q=2.62e-4$.

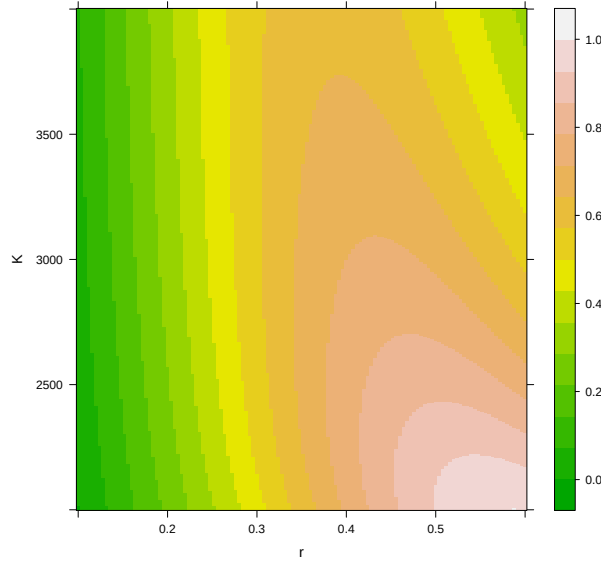


Figura 5.10: Curvas de nivel del logaritmo de la función de verosimilitud versus r y K para $q=6.34e-4$.

Para obtener los valores puntuales de los parámetros que optimizan la función de interés se tomó el mismo rango de valores iniciales para r y K que en el apartado anterior, además de que dadas las condiciones de este escenario, también se consideró un rango de valores iniciales para q , entre $1e-04$ y 1 , teniendo en cuenta el orden de valores obtenidos para este parámetro bajo el supuesto de incertidumbre en las observaciones.

La Tabla 5.4 contiene las medidas de resumen de los valores iniciales y los valores óptimos de los parámetros, así como del valor óptimo de la función de verosimilitud que se obtuvo para cada terna de valores iniciales de r , K y q .

Parámetro/Resumen	Min.	1 ^{er} Cuart.	Mediana	Media	3 ^{er} Cuart.	Máx.
r inicial	0.1000	0.3000	0.5000	0.5124	0.7500	1.0000
K inicial	2000	5000	7000	6557	9000	10000
q inicial	0.0001	0.0002	0.0003	0.0002735	0.0004	0.0004
r óptimo	0.1780	0.2860	0.3160	0.2951	0.3160	0.6470
K óptimo	1881	3516	3516	4404	3998	10333
q óptimo	6.422e-05	2.199e-04	2.624e-04	2.314e-04	2.624e-04	6.348e-04
$-sumLL$	-20.49	-20.49	-20.49	-20.21	-20.29	-15.38
RMS óptimo	253.0	278.0	278.0	300.8	286.0	460.0

Tabla 5.4: Estadística descriptiva del ajuste del modelo

A diferencia del escenario con incertidumbre en las observaciones, en este caso se observa variabilidad en los valores óptimos obtenidos para r , K y q y en el valor óptimo de la función de verosimilitud.

Valores iniciales			Óptimos						
r	K	q	r	K	σ_{proc}	q	RMS	$B_{\text{óptima}}$	$sum(-LL)$
0.1	5000	0.0001	0.316	3516	0.101	0.00026	278	1758	-20.4945

Tabla 5.5: Valores iniciales y valores óptimos

Como se mencionó anteriormente si bien se encontró mayor variabilidad en los valores que optimizan la función de interés que bajo el escenario de incertidumbre en las observaciones, el valor de $r = 0,316$ y $K = 3516$, es la combinación de parámetros con la cual se encuentra el mínimo valor de la función del logaritmo de la verosimilitud. Con los valores óptimos de r , K y q presentados en la Tabla 5.5, el ajuste del modelo de Schaefer para la variable CPUE y la Biomasa, bajo el supuesto de incertidumbre en el proceso, se presentan en las Figuras 5.11 y 5.12 respectivamente.

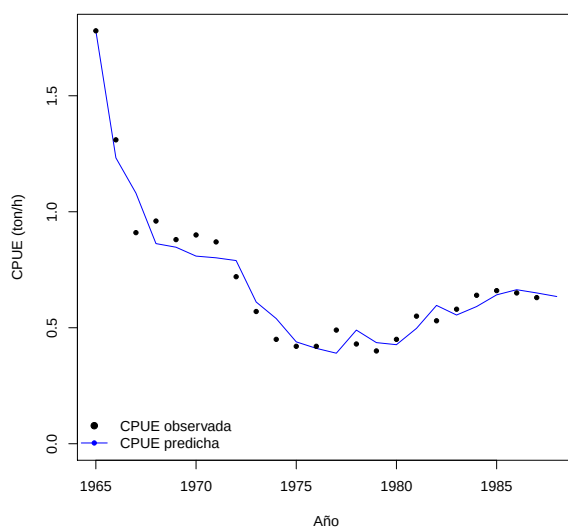


Figura 5.11: Evolución de la *CPUE* observada y la *CPUE* predicha por el modelo.

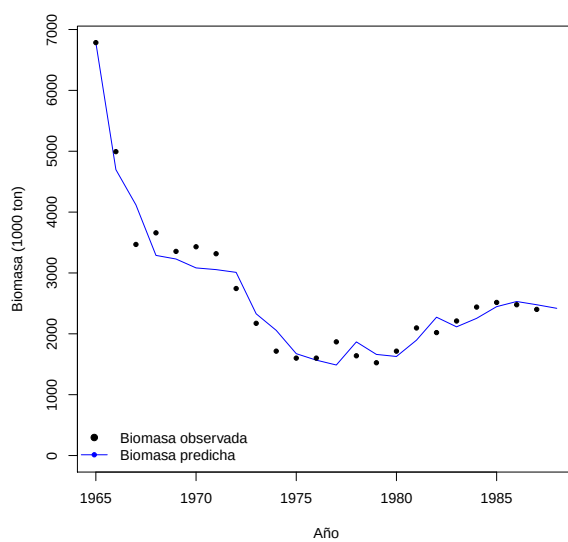


Figura 5.12: Evolución de la Biomasa observada y la Biomasa predicha por el modelo.

Resultados de las pruebas de normalidad

La Tabla 5.6 contiene el p-valor de cada una de las pruebas de normalidad realizadas para el residuo estimado del modelo de schaefer al 5 % de significación.

Residuo estimado/Prueba	Shapiro - Wilk	Kolmogoróv - Smirnov
$\ln(CPUE_{Obs.}) - \ln(CPUE_{Pred.})$	0.77	0.94

Tabla 5.6: Pruebas de normalidad

Las pruebas de normalidad indicadas en la Tabla 5.6 para el residuo del ajuste, dieron como resultado que no se rechaza la hipótesis de normalidad al 5 % de significación.

Perfiles de verosimilitud

Para realizar los perfiles de verosimilitud de r , K y del RMS , se debe evaluar para un rango de valores de cada parámetro la función de verosimilitud del mismo independientemente del valor que toman los restantes. A modo de ejemplo para un rango de valores de r se consideró un valor de K y además bajo este escenario también se consideró un valor de q y se aplicó el algoritmo de optimización de tal manera que para cada valor perteneciente al rango de r se optimizara la función del logaritmo de verosimilitud según los parámetros K y q . En las Figuras 5.13, 5.14 y 5.15 se presentan los perfiles de verosimilitud para los parámetros r y K y para el punto de referencia biológico RMS .

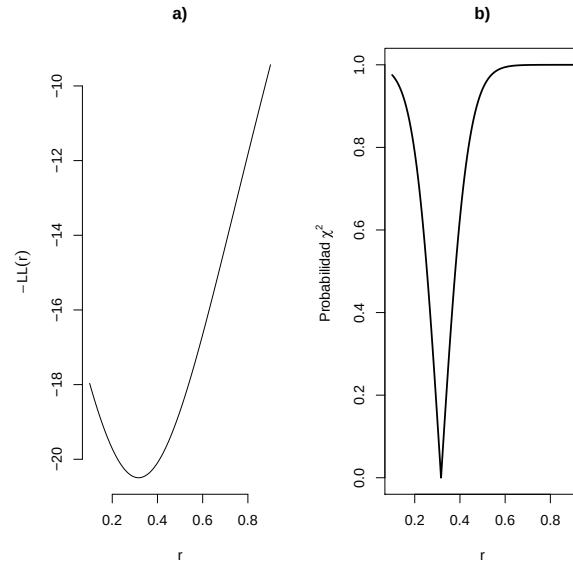


Figura 5.13: a) Perfil de verosimilitud de r ; b) Probabilidad Chi-cuadrado asociada a r .

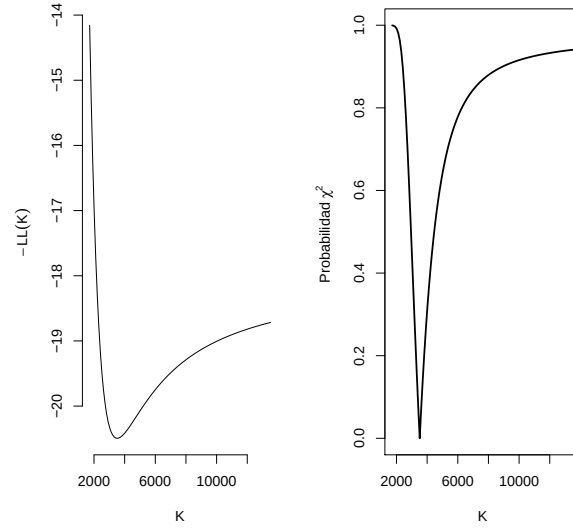


Figura 5.14: a) Perfil de verosimilitud de K ; b) Probabilidad Chi-cuadrado asociada a K .

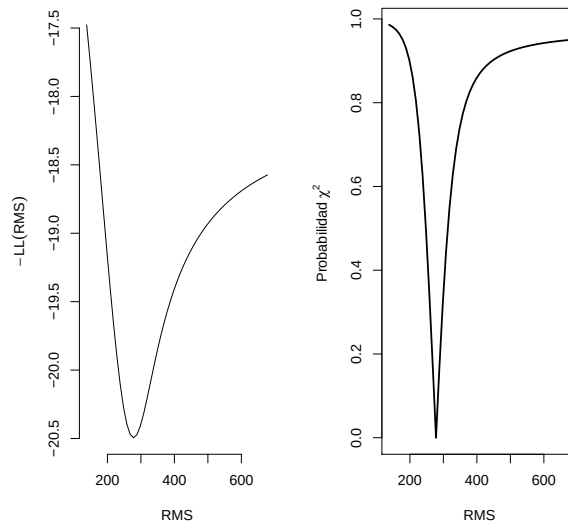


Figura 5.15: a) Perfil de verosimilitud de RMS ; b) Probabilidad Chi-cuadrado asociada al RMS .

Mediante las Figuras 5.13, 5.14, 5.15 y utilizando el criterio de la razón de verosimilitud planteado en el procedimiento, se obtienen los intervalos de confianza al 95 % para los parámetros de interés:

Parámetro	Límite inferior	Límite superior
r	0.13	0.51
K	2271	6000
RMS	180	600

5.1.4. Ajuste del modelo con el enfoque Bayesiano

Para realizar el ajuste del modelo por medio de inferencia Bayesiana se tiene en cuenta las siguientes distribuciones a priori para los parámetros:

$$\begin{aligned}
 r &\sim \text{dlnorm}(-1.2628643, 1.5192832)C(0.1, 0.8) \\
 K &\sim \text{dlnorm}(7.947476, 0.8112984)C(1000, 8000) \\
 iq &\sim \text{dgamma}(0.001, 0.001)C(1000, 10000) \\
 q &= 1/iq \\
 \text{isigma}_2 &\sim \text{dgamma}(1.708603, 0.008613854) \\
 \sigma_2 &= 1/\text{isigma}_2 \\
 \text{itau}_2 &\sim \text{dgamma}(1.708603, 0.008613854) \\
 \tau_2 &= 1/\text{itau}_2
 \end{aligned}$$

Para la elección de los parámetros de la distribución a priori de r se considera una Log-Normal con cuyos cuantiles al 10 % y 90 % toman los valores 0.1 y 0.8 respectivamente, para más detalle ver en los Anexos el procedimiento realizado para calcular estos parámetros en con el Programa R. Para el parámetro K se consideró una distribución a priori Lo-Normal cuyos cuantiles al 10 % y 90 % toman los valores 1000 y 8000 respectivamente. Para la distribución de q se consideró un distribución no informativa $1/iq$ donde iq tiene una distribución a priori Gamma, definida de esta forma [12] sólo se tiene que tener en cuenta el intervalo que puede tomar los valores de q , para este caso el intervalo considerado es: 0.0001 a 0.001. Para las distribuciones a priori de los desvíos tanto de las observaciones como del proceso se considera una InversaGamma cuyos cuantiles al 10 % y 90 % toman los valores 0.05 y 0.15 respectivamente.

Se detalla a continuación el procedimiento realizado en OpenBUGS para ajustar el modelo aplicando inferencia Bayesiana (el código se adjunta en los Anexos):

$$\begin{aligned}
 \text{Pmed}_1 &= 0 \\
 P_1 &\sim \text{dlnorm}(\text{Pmed}_1, \text{isigma}_2)C(0.001, 2.0) \\
 \left. \begin{aligned}
 \text{Pmed}_t &= \log(\max(P_{t-1} + r \cdot P_{t-1} \cdot (1 - P_{t-1}) - C_{t-1}/K, 0.001)) \\
 P_t &\sim \text{dlnorm}(\text{Pmed}_t, \text{isigma}_2)C(0.001, 2.0)
 \end{aligned} \right\} 2 \leq t \leq N \\
 \left. \begin{aligned}
 \text{CPUEmed}_t &= \log(q \cdot K \cdot P_t) \\
 \text{CPUE}_t &\sim \text{dlnorm}(\text{CPUEmed}_t, \text{itau}_2) \\
 \text{CPUEest}_t &= e^{\text{CPUEmed}_t}
 \end{aligned} \right\} 1 \leq t \leq N \quad \text{RMS} = r \cdot K/4
 \end{aligned}$$

Los resultados obtenidos son los siguientes:

Parámetro	Media	Desvío	Percentil 2.5	Mediana	Percentil 97.5
r	0.3741	0.1046	0.1809	0.3703	0.5937
K	3193	873.6	1998	3009	5400
q	4.486e-4	9.9e-5	2.723e-4	4.436e-4	6.599e-4
RMS	279.4	36.95	204.1	279.2	354.8

En el mismo se muestreó 300000 veces y se tomaron las últimas 100000 muestras de las tres cadenas que se corrieron. A continuación se detallan las Figuras (5.16, 5.17, 5.18, 5.19) que muestran las distribuciones a posteriores para los distintos parámetros, las Figuras de cuantiles, los autocorrelogramas y la convergencia de las cadenas para los distintos parámetros.

Parámetro r

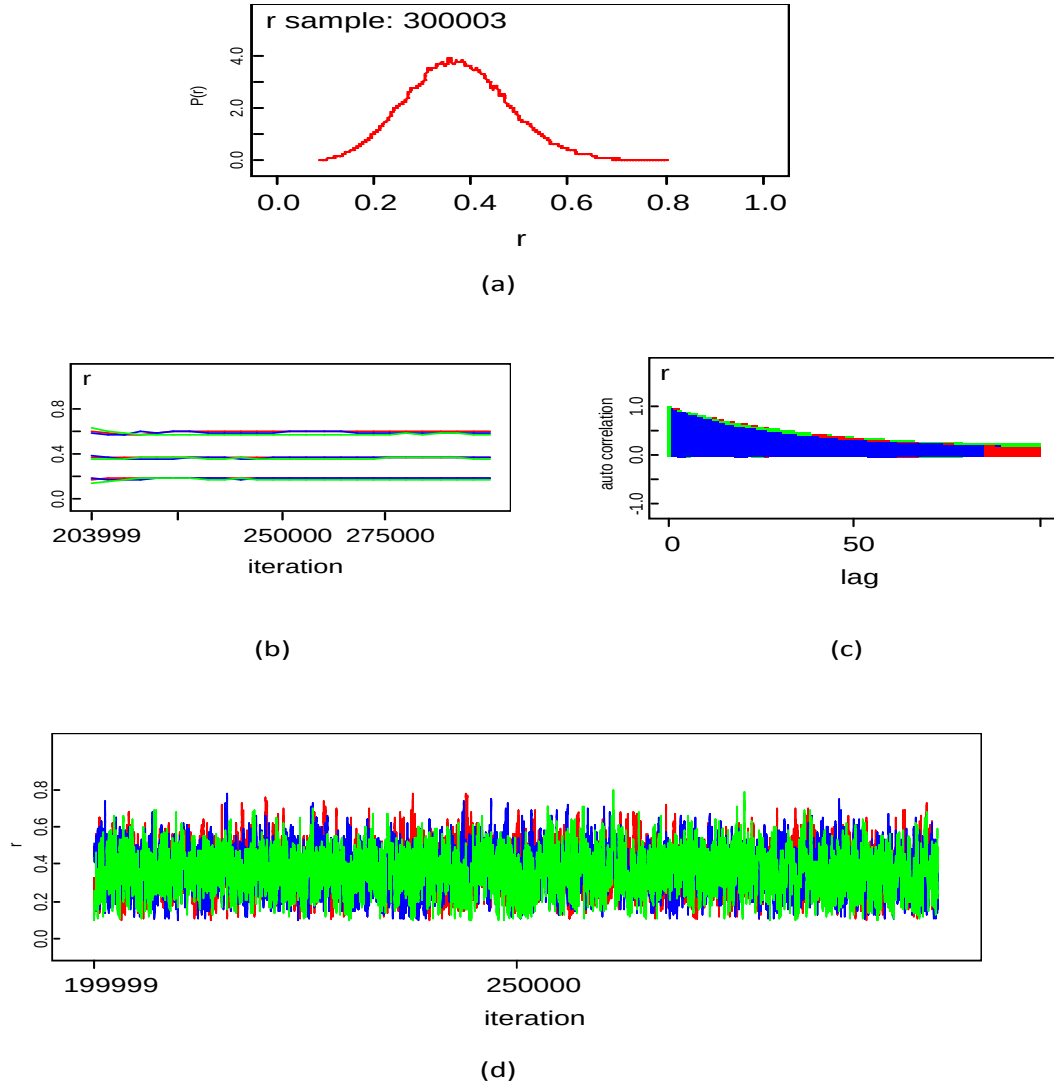


Figura 5.16: (a) Distribución a posteriori. (b) Cuantiles. (c) Autocorrelograma. (d) Historia que muestra la convergencia de las tres cadenas simuladas.

En la Figura 5.16(a) se observa una definida distribución a posteriori para el parámetro r centrada en 0.37. En la Figura 5.16(b) correspondiente a los cuantiles se observa

que los valores del parámetro están comprendidos en el intervalo 0.18 a 0.59. El autocorrelograma (5.16(c)) indica que no hay correlación con las iteraciones anteriores. En la evolución que indica la Figura (5.16(d)) se observa que hay convergencia en las tres cadenas simuladas para el parámetro r .

Parámetro K

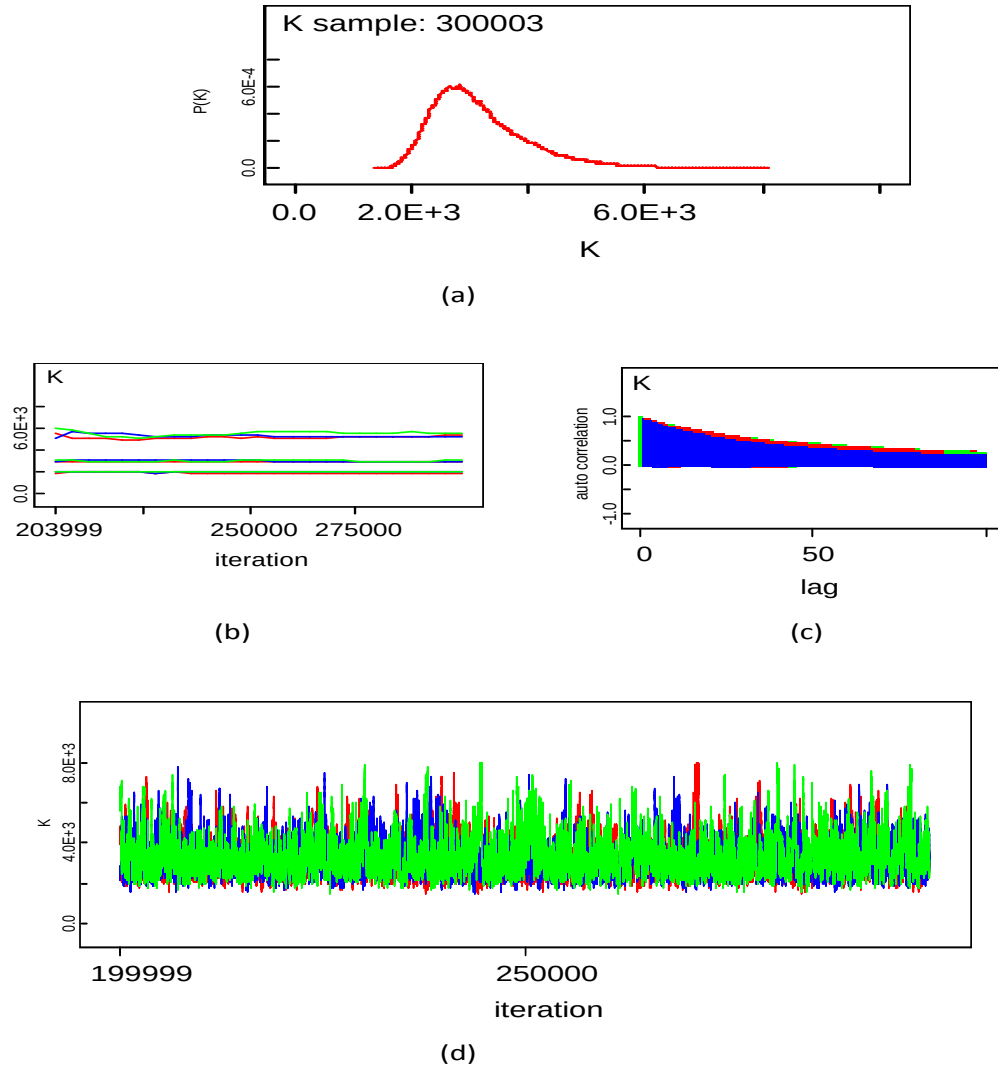


Figura 5.17: (a) Distribución a posteriori. (b) Cuantiles. (c) Autocorrelograma. (d) Historia que muestra la convergencia de las tres cadenas simuladas.

Al igual que para el parámetro r en la Figura 5.17(a) se observa una definida distribución a posteriori para el parámetro K con media 3193. En la Figura 5.17(b) correspon-

diente a los cuantiles se observa que los valores del parámetro para las tres cadenas están comprendidos en el intervalo 1998 a 5400. El autocorrelograma (5.17(c)) indica que no hay correlación con las iteraciones anteriores. La Figura (5.17(d)) indica que hay convergencia en las tres cadenas simuladas para el parámetro K .

Parámetro q

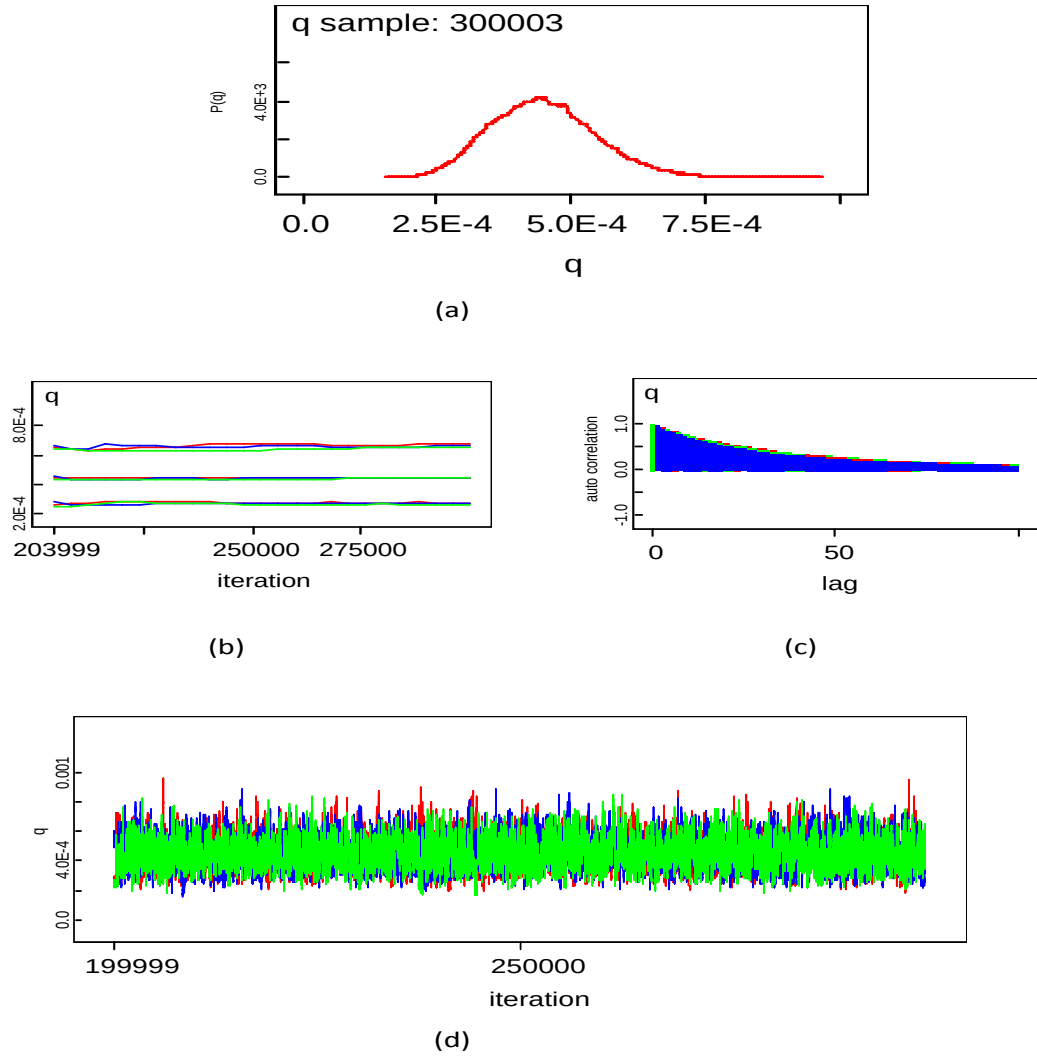


Figura 5.18: (a) Distribución a posteriori. (b) Cuantiles. (c) Autocorrelograma. (d) Historia que muestra la convergencia de las tres cadenas simuladas.

En la Figura 5.18(a) se observa una definida distribución a posteriori para q centrada en 4.48×10^{-4} . Los cuantiles de las tres cadenas simuladas (Figura 5.18(b)) indican que

los valores del parámetro q están comprendidos en el intervalo $2.72\text{e-}4$ a $6.59\text{e-}4$. El autocorrelograma (5.18(c)) indica que para los 100 lags no hay correlación con las iteraciones anteriores. En la evolución que indica la Figura 5.18(d) se observa que hay convergencia en las tres cadenas simuladas para el parámetro.

Parámetro RMS

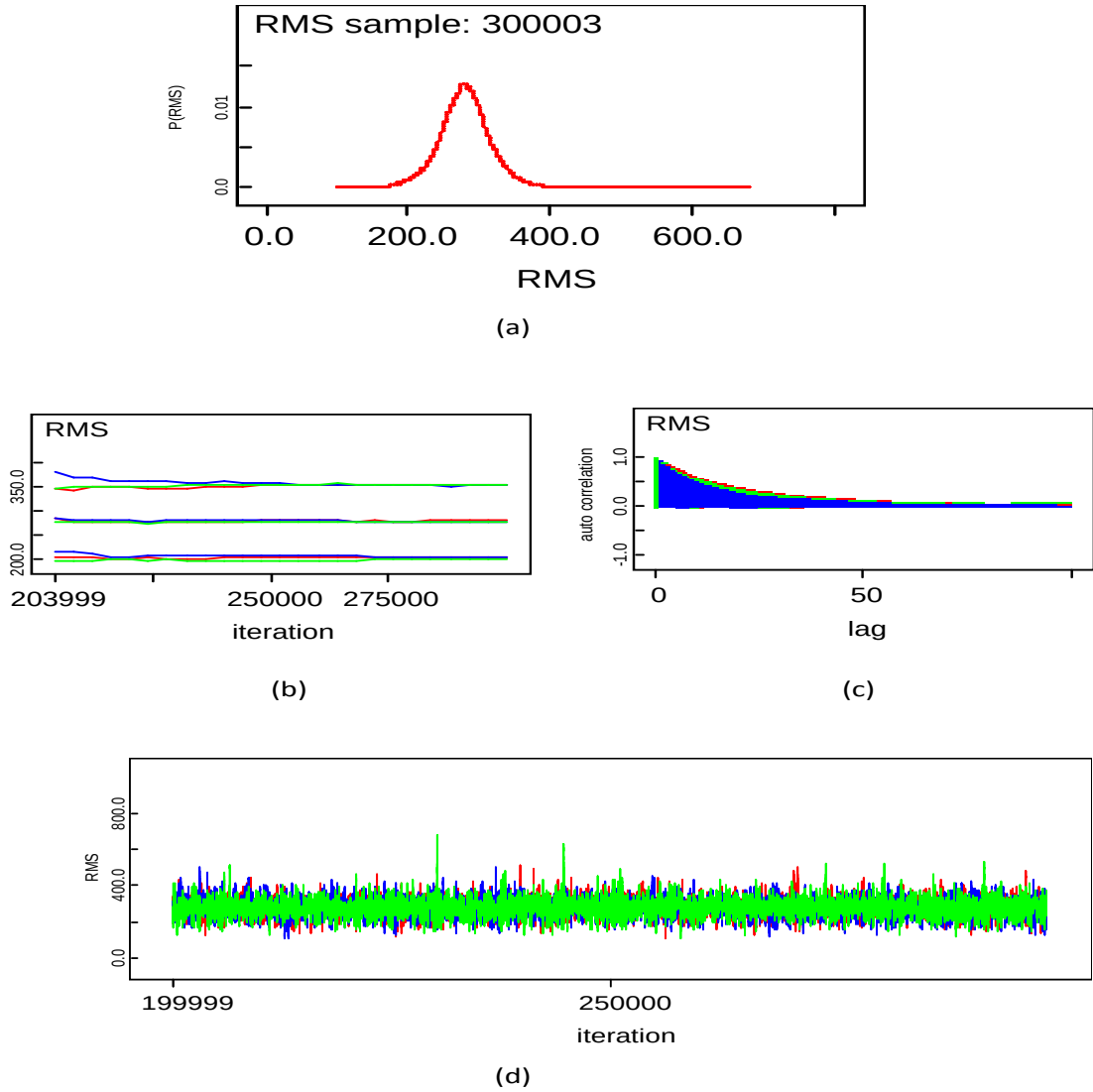


Figura 5.19: (a) Distribución a posteriori. (b) Cuantiles. (c) Autocorrelograma. (d) Historia que muestra la convergencia de las tres cadenas simuladas.

Por último para el parámetro RMS en la Figura 5.19(a) se observa una definida distribución a posteriori para ahora para el parámetro RMS con mediana 279. En la

Figura 5.19(b) correspondiente a los cuantiles se observa que los valores del parámetro para las tres cadenas están comprendidos en el intervalo 204 a 354. El autocorrelograma (5.19(c)) indica que no hay correlación con las iteraciones anteriores. La Figura 5.19(d) indica que hay convergencia en las tres cadenas simuladas para el parámetro RMS .

Comentarios finales sobre los resultados del enfoque Bayesiano

De los gráficos anteriores se desprende que hay convergencia en las tres cadenas simuladas para los cuatro parámetros de interés y los autocorrelogramas muestran una tendencia a disminuir con el paso del tiempo lo que indica que no hay correlación con el rezago anterior. Por ultimo en la Figura 5.20 se observa el ajuste del modelo con el valor medio de la $CPUE$.

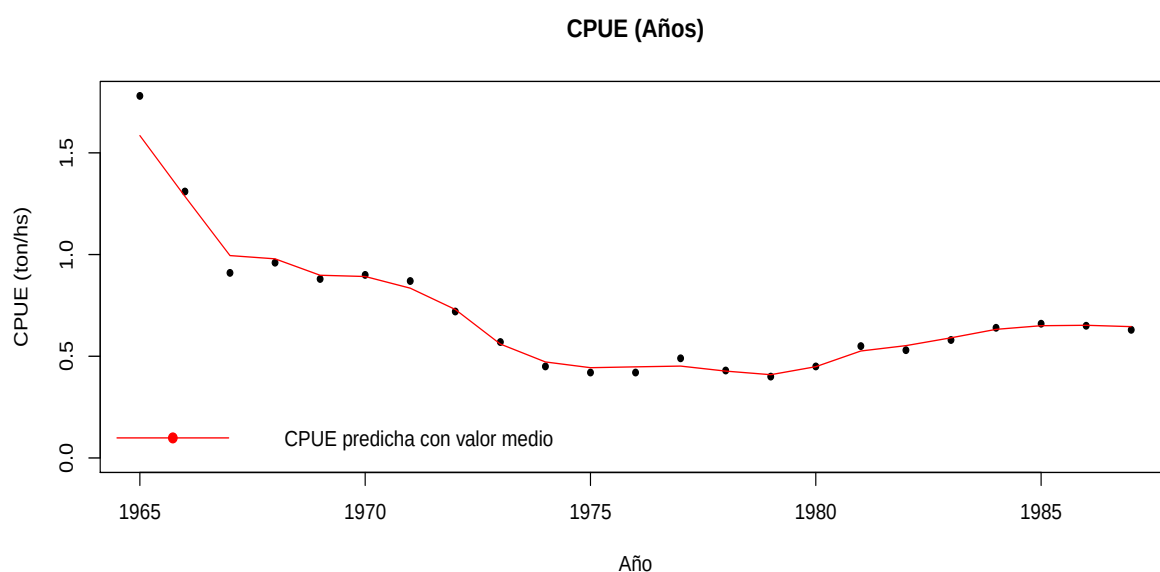


Figura 5.20: Ajuste del modelo de la variable $CPUE$ utilizando el valor medio de la salida de OpenBUGS.

La Figura 5.20 representa la serie de $CPUE$ observada y la evolución de la $CPUE$ trazada a partir de los valores medios generados por las iteraciones en el muestreo de

OpenBUGS.

5.2. Aplicación con datos de la Pesquería Uruguay

5.2.1. Unidades

Para los datos de la especie *Cynoscion guatucupa*, pescadilla, el parámetro K y la Biomasa se miden en toneladas, la tasa intrínseca de crecimiento r no tiene unidad, el coeficiente de capturabilidad q se mide en una proporción de 1/h y el Rendimiento Máximo Sostenible RMS se mide en toneladas.

5.2.2. Ajuste del modelo de Schaefer por Máxima Verosimilitud con incertidumbre en las observaciones

La pesquería de esta especie tiene sus inicios en el año 1975, dado que la serie de datos utilizada para el ajuste del modelo de Schaefer abarca el período 1996-2009, la suposición de que la Biomasa virgen, B_0 es igual a la capacidad de carga, K , no es aplicable en este caso. Por este motivo se considera un valor fijo para la Biomasa virgen menor o igual a la capacidad de carga (K). De acuerdo a la especie pescadilla, dichos valores de B_0 considerados son: 150000, 180000 y 210000 toneladas.

Gráficas de la función de verosimilitud

Para los distintos valores de B_0 mencionados anteriormente se realizaron las curvas de nivel para visualizar la o las regiones en donde los parámetros optimizan la función de interés normalizada.

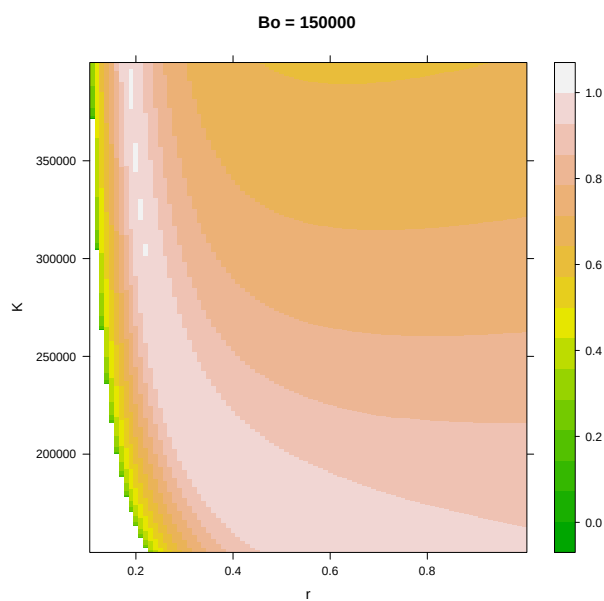


Figura 5.21: Superficie de nivel de la función de verosimilitud para $B_0 = 150000$.

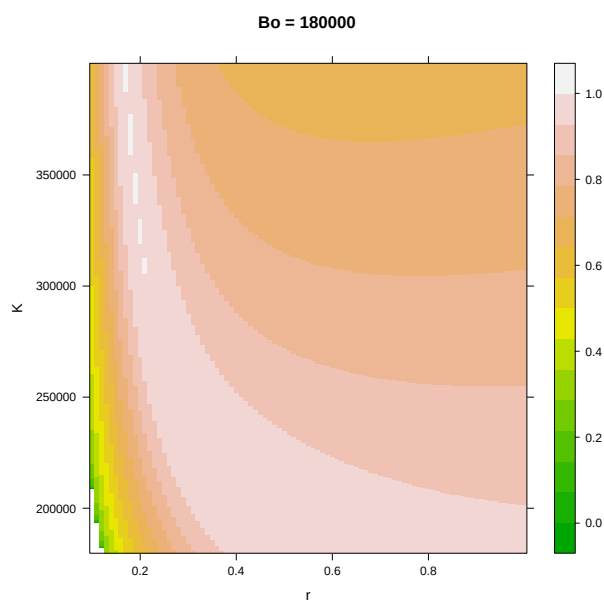


Figura 5.22: Superficie de nivel de la función de verosimilitud para $B_0 = 180000$.

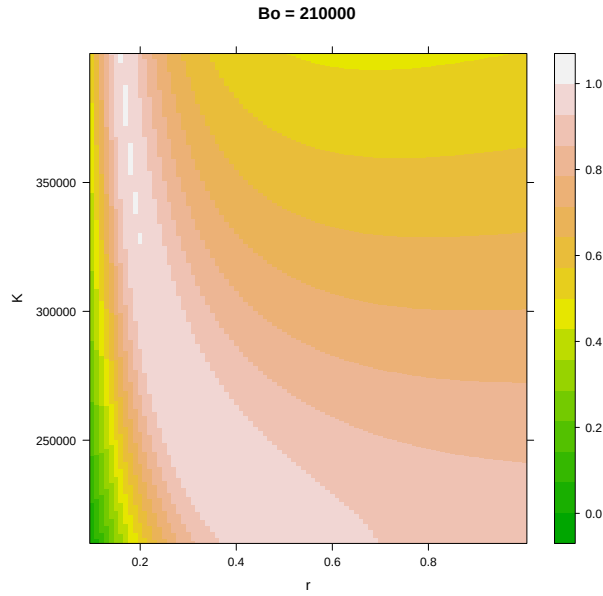


Figura 5.23: Superficie de nivel de la función de verosimilitud para $B_0 = 210000$.

Mediante el análisis de las Figuras anteriores se observa que en todo los casos existe múltiples combinaciones de los parámetros que minimizan la función de verosimilitud, o dicho de otra manera, no se encuentra una única región acotada de óptimos. Las Tablas 5.7, 5.8 y 5.9 contienen los resúmenes del mecanismo de optimización del modelo.

Parámetro/Resumen	Mín.	1 ^{er} Cuart.	Mediana	Media	3 ^{er} Cuart.	Máy.
r inicial	0.2000	0.4000	0.6000	0.6034	0.8000	1.0000
K inicial	150000	210000	280000	276034	340000	400000
r óptimo	0.1270	0.1370	0.1970	0.7789	1.7190	3.3790
K óptimo	152324	227503	384622	989324	1589842	8790784
$-sumLL$	-10.538	-10.532	-10.472	-8.859	-9.437	3.736
RMS óptimo	16302	32637	63050	65567	81213	279107

Tabla 5.7: Estadística descriptiva del ajuste del modelo para $B_0 = 150000$

Parámetro/Resumen	Mín.	1 ^{er} Cuart.	Mediana	Media	3 ^{er} Cuart.	Máy.
r inicial	0.1000	0.3000	0.5000	0.5459	0.8000	1.0000
K inicial	180000	230000	280000	286087	340000	400000
r óptimo	0.1070	0.1150	0.1260	0.4222	0.2265	3.3160
K óptimo	181842	308798	850094	929358	1429406	3206085
$-sumLL$	-10.5265	-10.5248	-10.5214	-9.9136	-10.4365	0.8809
RMS óptimo	16170	17125	33439	44653	48250	165279

Tabla 5.8: Estadística descriptiva del ajuste del modelo para $B_0 = 180000$

Parámetro/Resumen	Mín.	1 ^{er} Cuart.	Mediana	Media	3 ^{er} Cuart.	Máy.
r inicial	0.1000	0.3000	0.5000	0.5418	0.8000	1.0000
K inicial	210000	250000	300000	301693	350000	400000
r óptimo	0.0970	0.1040	0.1200	0.3444	0.2070	3.2880
K óptimo	210249	339556	668347	729319	1065553	1569611
$-sumLL$	-10.4655	-10.4651	-10.4630	-9.6486	-10.4132	-0.4417
RMS óptimo	16095	16950	25035	37312	31611	177460

Tabla 5.9: Estadística descriptiva del ajuste del modelo para $B_0 = 210000$

Para todos los caso de B_0 considerados se observa gran variabilidad en los valores óptimos de los parámetros, por lo tanto para elegir valores puntuales de los mismo se tiene en cuenta no sólo el valor que toma la función logaritmo de la verosimilitud, sino que también que el punto de referencia biológico, RMS , sea adecuado para la especie y el K máximo que se consideró como valor inicial. En la Tabla 5.10 se muestra algunos

de los valores óptimos, que minimizan la suma del logaritmo de la verosimilitud, de los parámetros y puntos de referencia biológicos adecuados para la pescadilla.

Valores iniciales			Óptimos						
B_0	r	K	r	K	σ_{obs}	q	RMS	$B_{\acute{o}ptima}$	$sum(-LL)$
150000	0.2	400000	0.187	400000	0.119	4.17e-06	18700	200000	-10.49
180000	0.2	400000	0.170	400000	0.120	3.46e-06	17000	200000	-10.50
210000	0.4	400000	0.162	399999	0.121	2.86e-06	16200	200000	-10.45

Tabla 5.10: Valores iniciales y valores óptimos según para cada valor fijo de biomasa virgen, B_0 .

Mediante los distintos valores óptimos presentados en la Tabla 5.10 se presentan a continuación gráficas con el ajuste, para cada valor de B_0 , de la $CPUE$ y de la Biomasa.

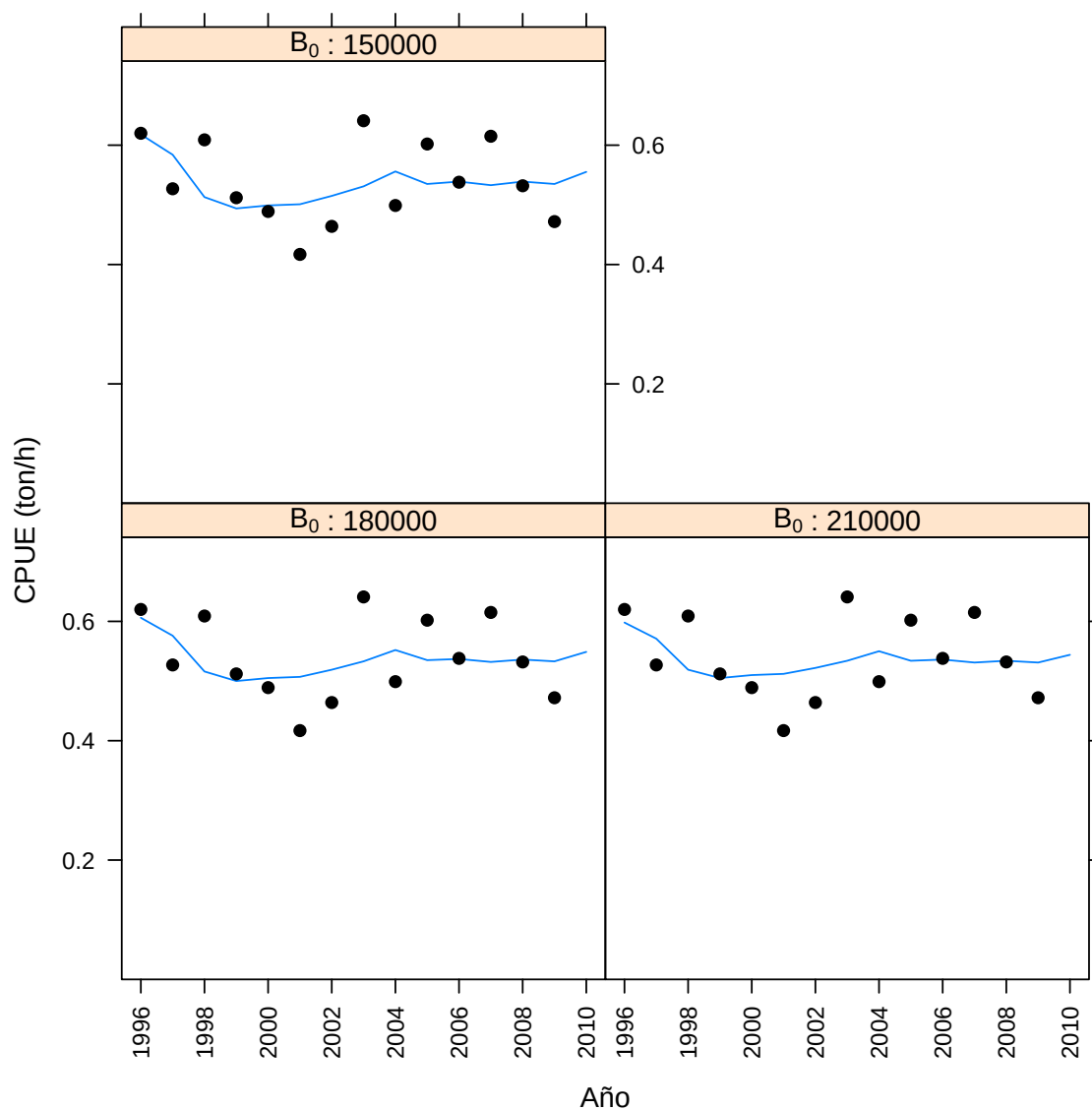


Figura 5.24: Evolución de la $CPUE$ observada(negro) y ajustada(azul) por el modelo para cada uno de los valores considerados de B_0 .

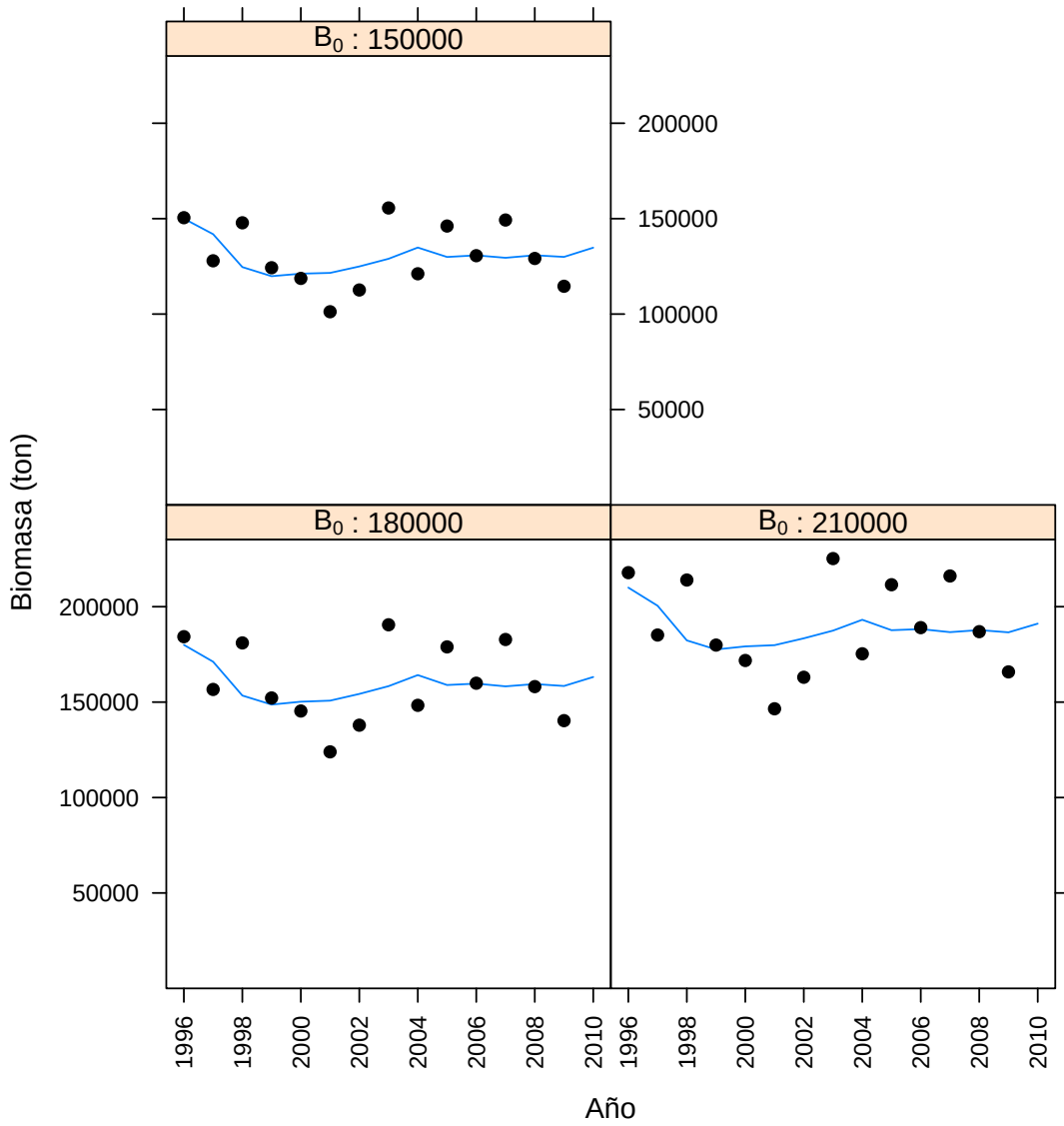


Figura 5.25: Evolución de la Biomasa observada(negro) y ajustada(azul) por el modelo para cada uno de los valores considerados de B_0 .

El resultado de la optimización con los datos de la especie pescadilla no es sensible a las condiciones iniciales, B_0 que se le imputan al modelo, a pesar de ello se obtienen valores de RMS biológicamente aceptables.

Resultados de las pruebas de normalidad

La Tabla 5.11 contiene para cada valor de B_0 , el p-valor de las pruebas de normalidad para la diferencia entre el logaritmo de la *CPUE* observada y el logaritmo de la *CPUE* ajustada por el modelo de Schaefer.

B_0	Shapiro - Wilk	Kolmogorov - Smirnov
150000	0.29	0.78
180000	0.40	0.94
210000	0.70	0.98

Tabla 5.11: Pruebas de normalidad del residuo estimado según los distintos valores de B_0

En todos los casos no se rechaza la hipótesis de normalidad de la diferencia entre el logaritmo de la *CPUE* observada y el logaritmo de la *CPUE* ajustada por el modelo de Schaefer, dado que el p-valor es mayor que el 5 % de significación.

Perfiles de verosimilitud

En las Figuras 5.26, 5.27 y 5.28 se presentan los perfiles de la verosimilitud de los parámetros r , K y RMS para los distintos casos considerados de B_0 .

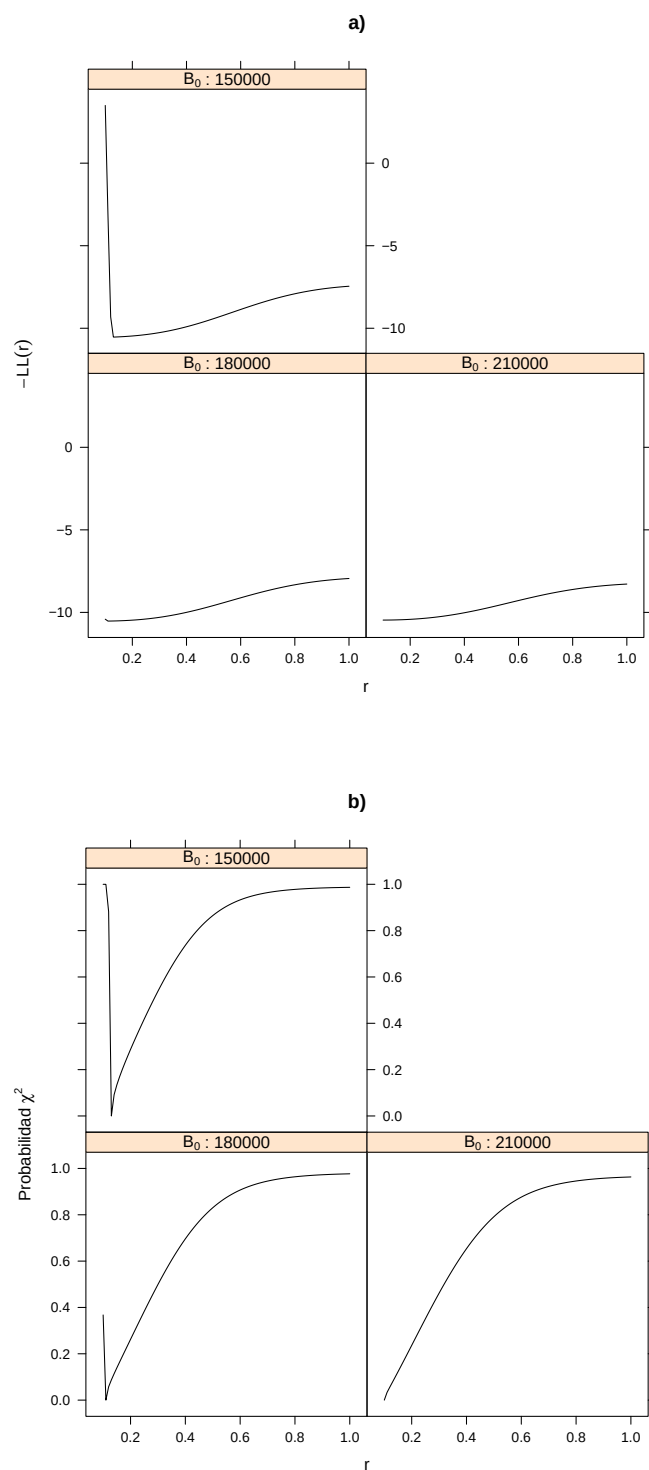


Figura 5.26: a) Perfil de verosimilitud de r ; b) Probabilidad Chi-cuadrado asociada a r según los valores de B_0 .

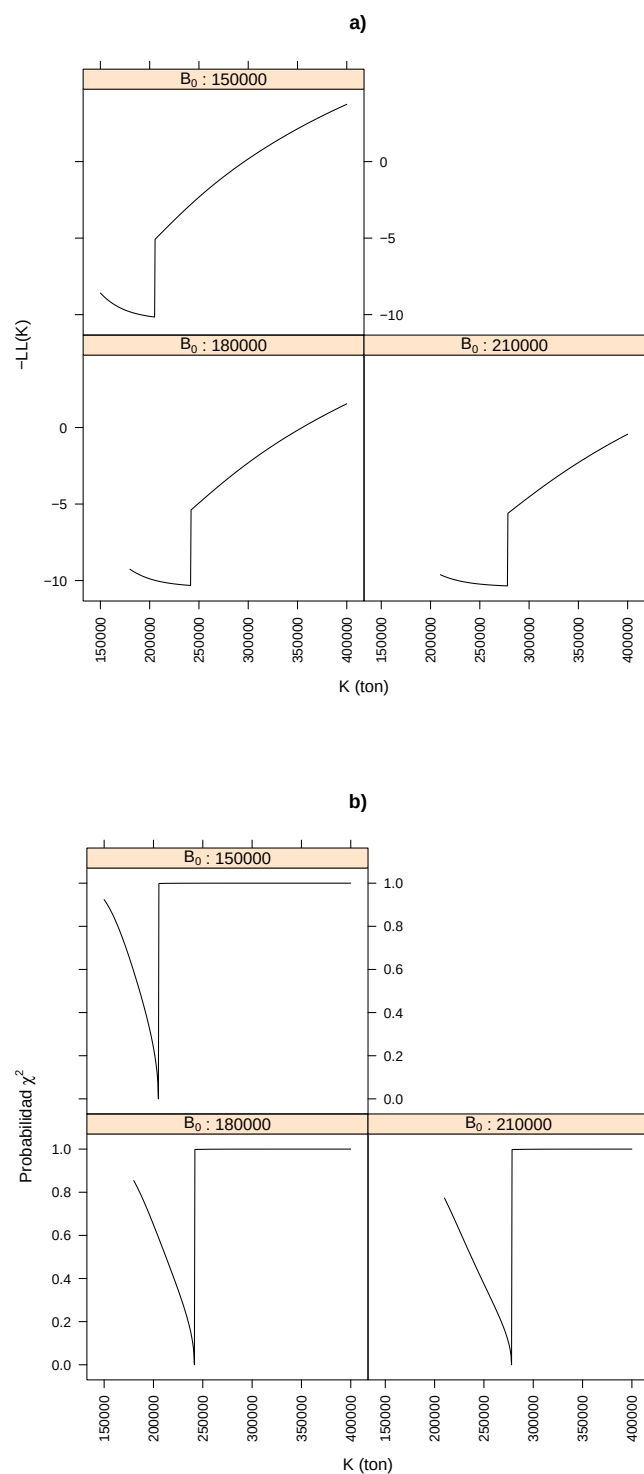


Figura 5.27: a) Perfil de verosimilitud de K ; b) Probabilidad Chi-cuadrado asociada a r según los valores de B_0 .

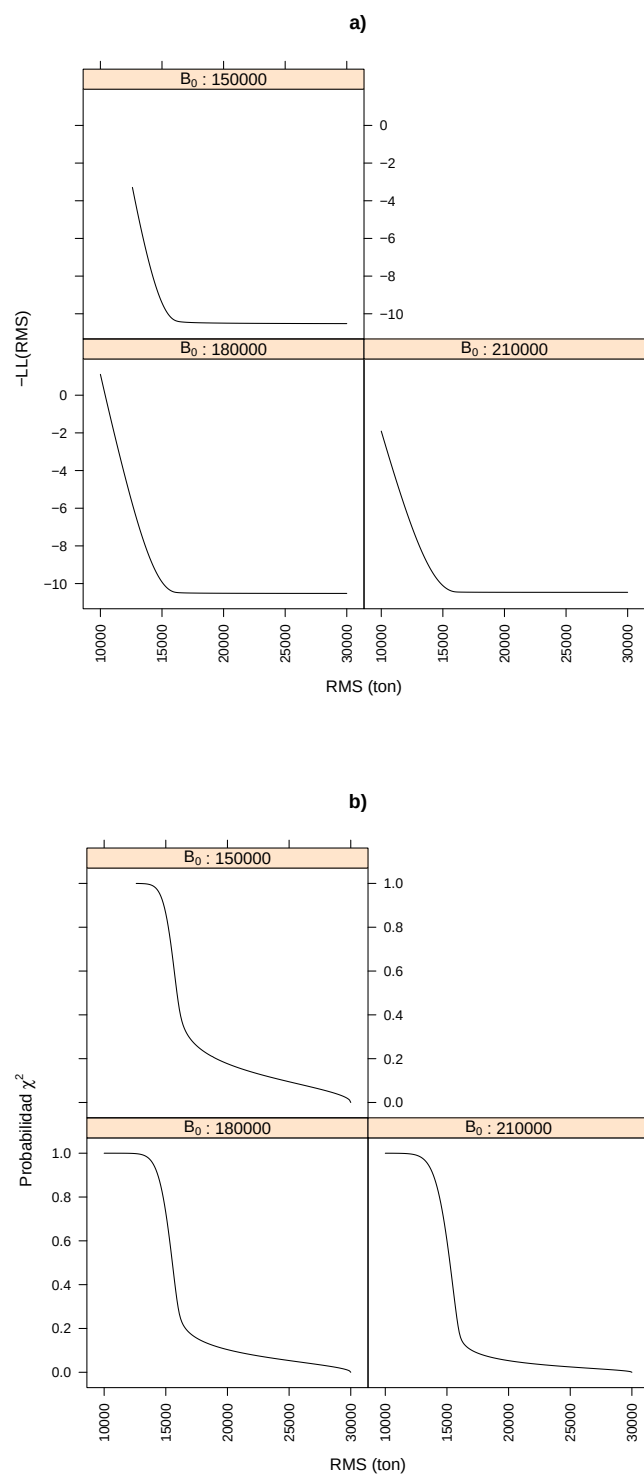


Figura 5.28: a) Perfil de verosimilitud de RMS ; b) Probabilidad Chi-cuadrado asociada al RMS según los valores de B_0 .

En este caso la metodología aplicada, si bien permite obtener valores de los parámetros r y K que minimizan la verosimilitud perfil, el mínimo que se obtiene para el punto de referencia RMS no es biológicamente aceptable para la especie y no coincide con los óptimos de la tabla 5.10.

5.2.3. Ajuste del modelo de Schaefer por Máxima Verosimilitud con incertidumbre en el proceso

El resultado del mecanismo de optimización para encontrar los parámetros del modelo de Schaefer bajo el supuesto de incertidumbre en el proceso se presentan en la Tabla 5.12

Parámetro/Resumen	Mín.	1 ^{er} Cuart.	Mediana	Media	3 ^{er} Cuart.	Máy.
r inicial	0.1000	0.3000	0.5000	0.5435	0.8000	1.0000
K inicial	150000	210000	280000	277147	340000	400000
q inicial	1.000e-06	4.000e-06	6.000e-06	5.865e-06	8.000e-06	1.000e-05
r óptimo	0.411	1.016	1.112	1.099	1.209	1.603
K óptimo	149886	290001	446797	922933	1126843	9803516
q óptimo	5.164e-08	4.827e-07	1.255e-06	1.343e-06	1.967e-06	4.018e-06
$-sumLL$	-10.394	-10.018	-9.614	-9.538	-9.177	-3.151
RMS óptimo	36185	84274	143955	208935	288080	1176522

Tabla 5.12: Estadística descriptiva del ajuste del modelo

El mínimo valor óptimo de rendimiento máximo sostenible (RMS óptimo) encontrado no es biológicamente aceptable para la especie en cuestión, por lo tanto este escenario no arroja resultados factibles como los obtenidos bajo el supuesto de incertidumbre en las observaciones.

5.2.4. Ajuste del modelo con el enfoque Bayesiano

Otra alternativa para ajustar el modelo de Schaefer es utilizar un enfoque Bayesiano, para obtener mayor información sobre los parámetros de interés ([11] Pág. 66). Para

realizar el ajuste del modelo por medio de inferencia Bayesiana se tiene en cuenta las siguientes distribuciones a priori para los parámetros:

$$\begin{aligned}
 r &\sim \text{dlnorm}(-1.497866, 2.536201)C(0.1, 0.6) \\
 K &\sim \text{dlnorm}(12.40881, 6.828815)C(150000, 400000) \\
 iq &\sim \text{dgamma}(0.001, 0.001)C(1000000, 10000000) \\
 q &= 1/iq \\
 \text{isigma}_2 &\sim \text{dgamma}(1.708603, 0.008613854) \\
 \sigma_2 &= 1/\text{isigma}_2 \\
 \text{itau}_2 &\sim \text{dgamma}(1.708603, 0.008613854) \\
 \tau_2 &= 1/\text{itau}_2
 \end{aligned}$$

Para la elección de los parámetros de la distribución a priori de r se considera una Log-Normal con cuyos cuantiles al 10 % y 90 % toman los valores 0.1 y 0.6 respectivamente, ver en los Anexos el procedimiento realizado para calcular estos parametros en con el Programa R. Para el parámetro K se consideró una distribución a priori Log-Normal cuyos cuantiles al 10 % y 90 % toman los valores 150000 y 400000 respectivamente. Para la distribución de q se consideró un distribucioón no informativa $1/iq$ donde iq tiene una distribución a priori Gamma, definida de esta forma [12] sólo se tiene que tener en cuenta el intervalo que puede tomar los valores de q , para este caso el intervalo considerado es: 1E-7 a 1E-6. Para las distribuciones a priori de los desvíos tanto de las observaciones como del proceso se considera una InversaGamma cuyos cuantiles al 10 % y 90 % toman los valores 0.05 y 0.15 respectivamente.

Se detalla a continuación el procedimiento realizado en OpenBUGS para ajustar el modelo aplicando inferencia Bayesiana:

$$\begin{aligned}
 \text{Pmed}_1 &= 0 \\
 P_1 &\sim \text{dlnorm}(\text{Pmed}_1, \text{isigma}_2)C(0.001, 2.0) \\
 \left. \begin{aligned}
 \text{Pmed}_t &= \log(\max(P_{t-1} + r \cdot P_{t-1} \cdot (1 - P_{t-1}) - C_{t-1}/K, 0.001)) \\
 P_t &\sim \text{dlnorm}(\text{Pmed}_t, \text{isigma}_2)C(0.001, 2.0)
 \end{aligned} \right\} 2 \leq t \leq N
 \end{aligned}$$

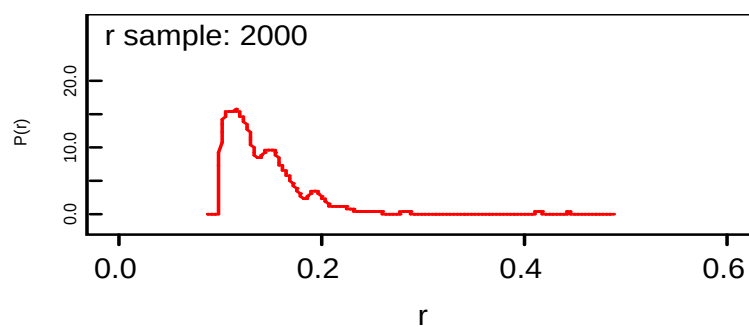
$$\left. \begin{aligned} \text{CPUEmed}_t &= \log(q \cdot K \cdot P_t) \\ \text{CPUE}_t &\sim \text{dlnorm}(\text{CPUEmed}_t, \text{itau}_2) \\ \text{CPUEest}_t &= e^{\text{CPUEmed}_t} \end{aligned} \right\} 1 \leq t \leq N \quad \text{RMS} = r \cdot K/4$$

Los resultados obtenidos son los siguientes:

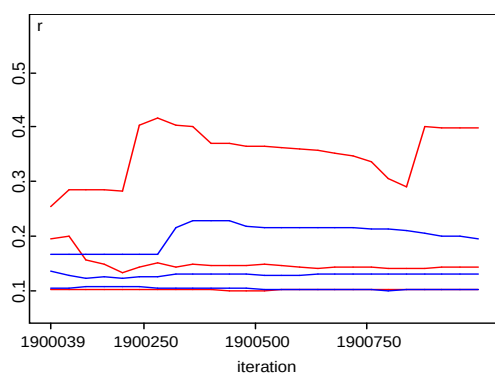
Parámetro	Media	Desvío	Percentil 2.5	Mediana	Percentil 97.5
r	0.1476	0.05112	0.1012	0.1339	0.2849
K	378900	17990	333600	384000	399400
q	9.99e-7	9,29e-7	9,97e-4	9,99e-4	1.0e-6
RMS	14010	5094	9093	12670	27770

En el mismo se muestreó 2000000 veces para que las cadenas converjan y se tomaron las últimas 100000 muestras de las dos cadenas que se corrieron. A continuación se detallan gráficamente las distribuciones a posteriori para los distintos parámetros, las gráficas de cuantiles, los autocorrelogramas y la convergencia de las cadenas para los distintos parámetros (Figuras 5.29, 5.30, 5.31, 5.32).

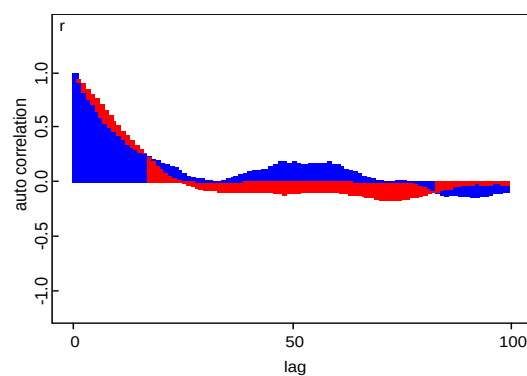
Parámetro r



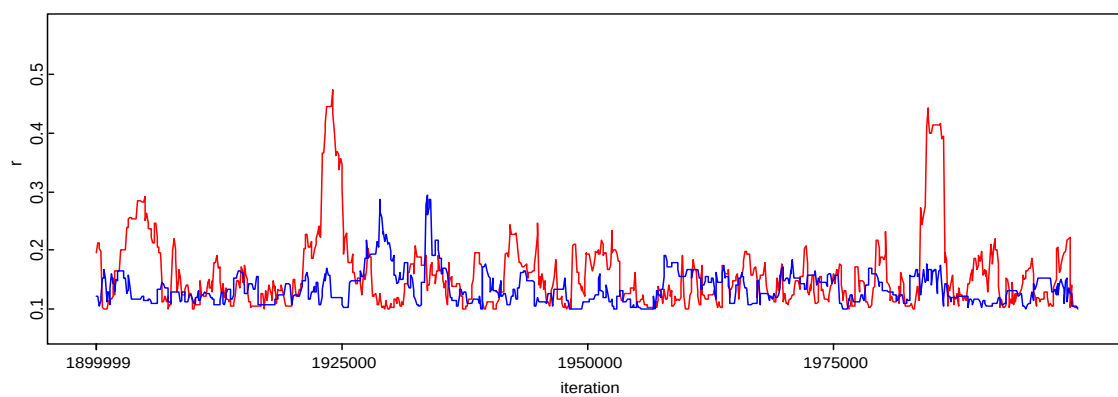
(a)



(b)



(c)



(d)

Figura 5.29: (a) Distribución a posteriori. (b) Cuantiles. (c) Autocorrelograma. (d) Historia que muestra la convergencia de las tres cadenas simuladas.

En la Figura 5.29(a) se observa la distribución a posteriori para el parámetro r , la curva de la distribución no esta bien definida como ocurriría en la Figura 5.16(a). En la Figura 5.29(b) correspondiente a los cuantiles se observa que una de las dos cadenas no define un extremos lineal del cuantil. El autocorrelograma (5.29(c)) indica que no hay correlación con las iteraciones anteriores. En la evolución que indica la Figura (5.16(d)) se observa que hay convergencia en las dos cadenas simuladas para el parámetro r .

Parámetro K

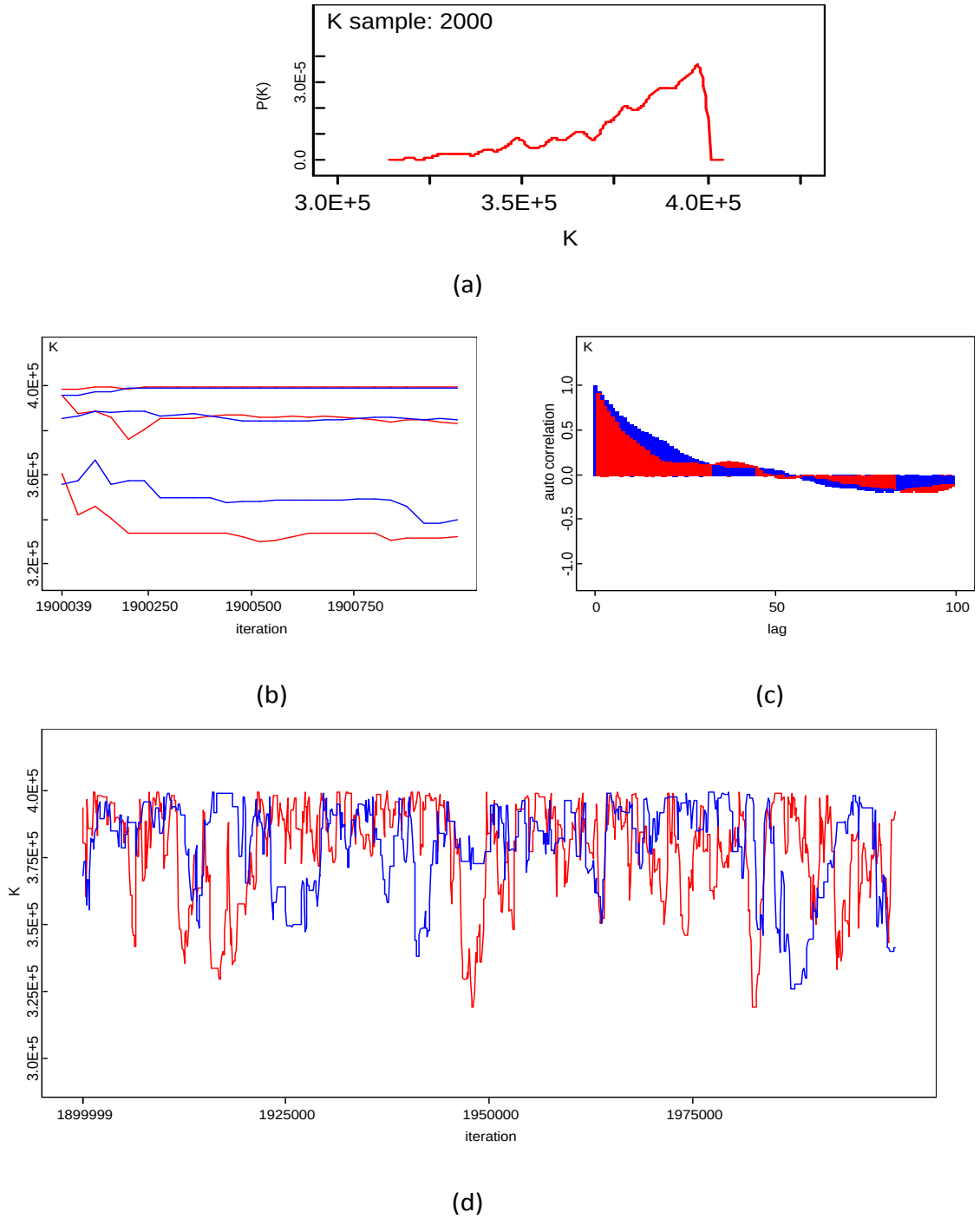


Figura 5.30: (a) Distribución a posteriori. (b) Cuantiles. (c) Autocorrelograma. (d) Historia que muestra la convergencia de las tres cadenas simuladas.

Al igual que para el parámetro r en la Figura 5.30(a) no se observa una curva definida para la distribución a posteriori del parámetro K y la distribución parece estar concentrada en el extremo de 400000. En la Figura 5.30(b) correspondiente a los cuantiles se observa que los valores del parámetro para las dos cadenas están comprendidos en el intervalo 320000 a 400000. El autocorrelograma (5.30(c)) indica que no hay correlación con las iteraciones anteriores. La Figura (5.30(d)) indica que hay convergencia en las dos cadenas simuladas para el parámetro K .

Parámetro q

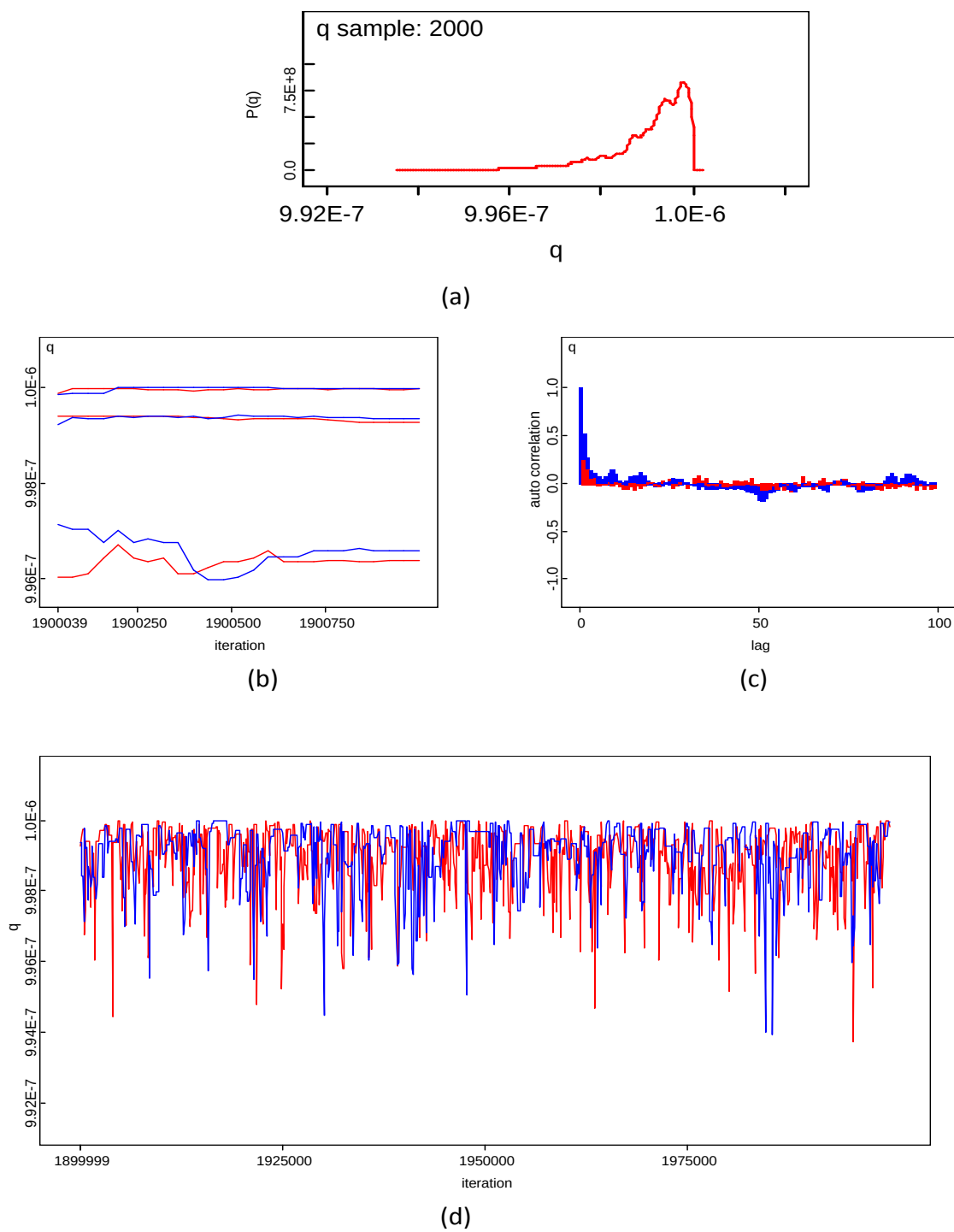


Figura 5.31: (a) Distribución a posteriori. (b) Cuantiles. (c) Autocorrelograma. (d) Historia que muestra la convergencia de las tres cadenas simuladas.

En la Figura 5.31(a) no se observa una curva definida en la distribución a posteriori del parámetro q y la misma está concentrada en uno de sus extremos. Los cuantiles de las dos cadenas simuladas (Figura 5.31(b)) indican que los valores del parámetro q están comprendidos en el intervalo $0.96e-7$ a $1e-7$. El autocorrelograma (5.31(c)) indica que no hay correlación con las iteraciones anteriores. En la evolución que indica la Figura 5.31(d) se observa que hay convergencia en las dos cadenas simuladas para el parámetro.

Parámetro RMS

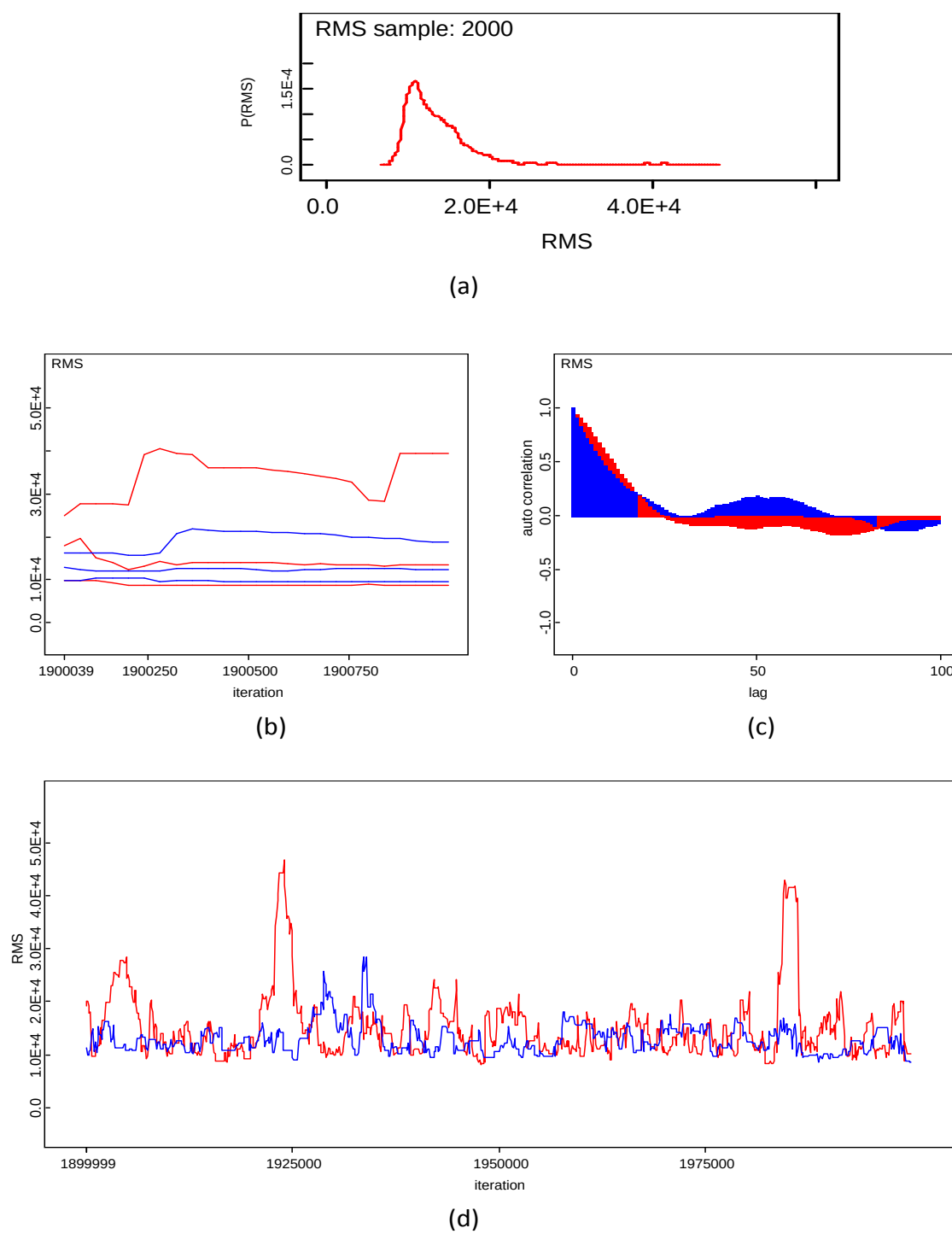


Figura 5.32: (a) Distribución a posteriori. (b) Cuantiles. (c) Autocorrelograma. (d) Historia que muestra la convergencia de las tres cadenas simuladas.

Por último para el parámetro RMS en la Figura 5.32(a) se observa una definida distribución a posteriori. En la Figura 5.32(b) correspondiente a los cuantiles se observa que una de las dos cadenas no define un extremo lineal del cuantil. El autocorrelograma (5.32(c)) indica que no hay correlación con las iteraciones anteriores. La Figura 5.32(d) indica que hay convergencia en las dos cadenas simuladas para el parámetro RMS .

Comentarios finales sobre los resultados del enfoque Bayesiano

De los gráficos anteriores se desprende que hay convergencia en las dos cadenas simuladas para los cuatro parámetros de interés, esto se logró después de simular muchas veces la convergencia se obtenía cerca de la simulación 300000. Por otro lado los autocorrelogramas muestran una tendencia a disminuir con el paso del tiempo lo que indica que no hay correlación con el rezago anterior. Por ultimo se adjunta la Figura 5.33 del ajuste del modelo con el valor medio de la $CPUE$.

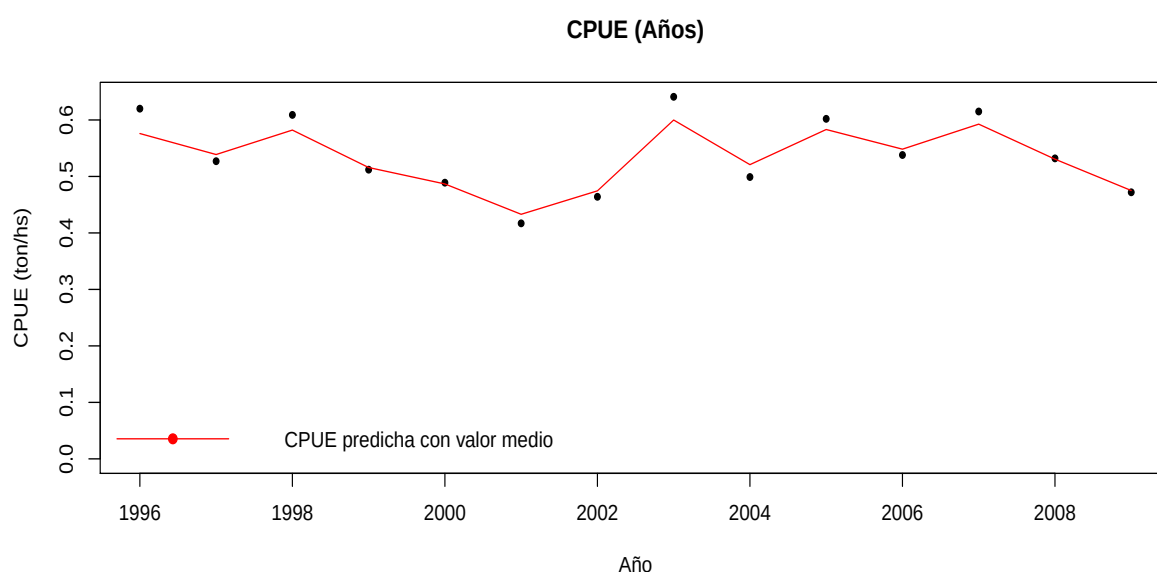


Figura 5.33: Ajuste del modelo de la variable $CPUE$ utilizando el valor medio de la salida de OpenBUGS.

Capítulo 6

Conclusiones

De lo anteriormente expuesto se pueden destacar algunas conclusiones como las que se detallan a continuación:

1. Es posible generar a través del lenguaje en programación R, un procedimiento que permita encontrar los óptimos de los parámetros por máxima verosimilitud, de la tasa intrínseca de crecimiento y la capacidad de carga con la ventaja de que permite tener estimaciones puntuales con intervalos de confianza y genera gráficos de superficies de nivel que permiten encontrar regiones de optimización. Por otro lado el procedimiento permite construir perfiles de verosimilitud que permiten elegir los mejores óptimos de los estimadores de interés.
2. A través del procedimiento desarrollado se logró ajustar modelos estocásticos sobre procesos de dinámica poblacional y por consiguiente esto permite evaluar las implicancias de manejo alternativas sobre recursos pesqueros por medio de los valores en los puntos de referencia biológica.
3. Los análisis de sensibilidad estadística al igual que las pruebas de normalidad realizados permiten concluir que el procedimiento es confiable, dado que en los mismos se obtuvieron resultados satisfactorios.
4. Una de las limitaciones que se ha encontrado, es que el procedimiento de ajuste del

modelo Schaefer por máxima verosimilitud, es débil cuando las series cronológicas de las variables *CPUE* y *Captura* tienen ciertas características, como es el caso de los datos de la especie *Cynoscion guatucupa*, pescadilla, en este caso el modelo Schaefer no ajusta de la mejor forma, pero se ha optado por tener valores de *RMS* biológicamente correctos.

5. Una alternativa para obtener mejor información sobre los parámetros del modelo de Schaefer es aplicar un método Bayesiano. El cual permite encontrar distribuciones a posteriori de los parámetros. Por otro lado las estimaciones puntuales de los parámetros realizadas por máxima verosimilitud se encuentran contempladas dentro de los posibles valores de las distribuciones a posteriori de los parámetros.

Capítulo 7

Sugerencias para trabajos futuros

De los posibles aportes para futuros trabajos se pueden destacar los siguientes:

- Implementar el algoritmo creado para otros modelos dinámicos de producción como por ejemplo el modelo de Fox.
- Implementar los procedimientos realizados en *modelos estructurados* para la evaluación.
- Probar otras formas de incorporar la incertidumbre en las observaciones y/o en el proceso.
- Lograr trabajar con series cronológicas de las variables *CPUE* y *Captura* que contengan mayor cantidad de datos.
- Realizar el ajuste del modelo de Schaefer por máxima verosimilitud a los datos de la Pesquería Uruguaya, utilizando otras formas alternativas para la *Tasa intrínseca de crecimiento* y la *Capacidad de carga*, por ejemplo que las mismas dependan del tiempo.
- Utilizar otras formas de optimización en R, como por ejemplo usar el Paquete "DEoptim".
- En el enfoque Bayesiano utilizar otras distribuciones a priori para los parámetros.

Bibliografía

- [1] Arena G., Gamarra M., “Captura Máxima Sostenible de Pescadilla”, “Modelos de Producción Excedente Aplicados a los Recursos Corvina y Pescadilla”, Proyecto URU /92/003, INAPE - PNUD, Pág. 80 a 103.
- [2] Bacaër N. “A Short History of Mathematical Population Dynamics”, Springer, 2011.
- [3] Brauer F., Castillo-Chávez C. “Mathematical Models in Population Biology and Epidemiology”, Springer, 2001.
- [4] Cadima E. L. , “Manual de evaluación de recursos pesqueros, Organización de las Naciones Unidas para la Agricultura, FAO, Documento de Pesca 339, Roma, 2003".
- [5] Cooke K., Wittens M., “One Dimensional Linear and Logistic Harvesting Models”, Mathematical Modelling, Vol. 7, pp. 301-340, Pergamon Journals Ltd., U.S.A., 1986.
- [6] Del Monte Luna, P., “Confrontación de modelos de la Biomasa Dinámica: una aplicación a la Pesquería del Huachinango”. Instituto Politécnico Nacional. La Paz, Baja California Sur, México, 2010.
- [7] Gelman A., Carlin J.B., Stern H.S., Rubin D.B.,”Bayesian Data Analysis”, Chapman and Hall/CRC, U.S.A., 2004.
- [8] Haddon M., “Modelling and Quantitative Methods in Fisheries”, Chapman and Hall/CRC, 2001.

- [9] Hilborn R., Mangel M., “The Ecological Detective: Confronting Models with Data”, 1997.
- [10] May R. M., Conway G. R., Hassell M. P., Southwood T. R. E., “Time Delays, Density-Dependence and Single-Species Oscillations”, *Journal of Animal Ecology*, Vol. 43, No. 3 (Oct., 1974), Págs. 747-770.
- [11] McLeod K., “Risk analysis of a Flatfish stock complex”, Thesis presented in partial fulfilment of the requirements for the degree of Master of Science in Mathematics at Massey University, New Zealand, 2010.
- [12] Meyer R., Millar R.B., “Bugs in Bayesian stock assessments”, *Canadian Journal of Fisheries and Aquatic Sciences* 56, Pág.: 1078 a 1086, Canada, 1999.
- [13] Punt A.E., Hilborn R., “Fisheries stock assessment and decision analysis: the Bayesian approach”. *Reviews in Fish Biology and Fisheries* 7, Págs. 35 a 63, 1997.
- [14] Sparre P., Venema Siebren C. “Introducción a la evaluación de recursos pesqueros tropicales, Organización de las Naciones Unidas para la Agricultura, FAO, Documento Técnica de Pesca, 306/1, Rev.2, Roma 1997”.
- [15] Thode H. C., “Testing for Normality”, *Statistics: textbooks and monographs*, Volume 164, Cap. 2 y Cap. 5, Marcel Dekker Inc., 2002.
- [16] Paquete estadístico utilizado: R Development Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
- [17] Paquete de Análisis Bayesiano utilizado: OpenBUGS 3.2.1, <http://www.openbugs.info>.

Capítulo 8

Anexo

8.1. Script de R para realizar el ajuste por máxima verosimilitud

El presente script de R para ajustar el modelo por Máxima Verosimilitud, fue realizado para que lo utilicen los investigadores de la DINARA. El mismo utiliza varias funciones con el fin de facilitar el uso del algoritmo.

8.1.1. Funciones

```
rangoptim <- function(K.inicial, K.final, K.intervalo,
  r.inicial, r.final, r.intervalo, q.inicial = NULL,
  q.final = NULL, q.intervalo = NULL, Bo = NULL)
{
  n = nrow(datos) #cantidad de años
  K = seq(K.inicial, K.final, K.intervalo)
  r = seq(r.inicial,r.final,r.intervalo)
```

```
# Error en las observaciones
if(is.null(q.inicial))
{
  aux = expand.grid(r = r, K = K)
  if(!is.null(Bo)){aux = aux[aux$K >= Bo,]}
  aux$r.optimo = rep(0, nrow(aux))
  aux$K.optimo = rep(0, nrow(aux))
  aux$sumLL = rep(0, nrow(aux))
  for (j in 1:nrow(aux)){
    print(j)
    loglik2011 <- function(x){
      aux$r[j] = x[1]
      aux$K[j] = x[2]
      Bpred <- numeric(n) #Biomasa predicha por el modelo de Schaefer
      if(is.null(Bo)){Bpred[1] = aux$K[j]}
      if(!is.null(Bo)){Bpred[1] = Bo}
      for (i in 2:n){
        Bpred[i] = Bpred[i-1]+((aux$r[j]*Bpred[i-1])*(1-
          (Bpred[i-1]/aux$K[j])))-datos[i-1,3] #modelo de Schaefer
      }
      Log <-(log(datos[,2]/Bpred))
      q <- exp((1/n)*sum(Log)) #Coeficiente de capturabilidad
      Bobs = datos[,2]/q #Biomasa observada
      CPUEpred = Bpred*q #CPUE predicha por el modelo
      sd.obs = sqrt(sum(((log(datos[,2])-log(CPUEpred))^2)/(n-1)))
      Lt = sum(-dnorm(log(datos[,2]),log(CPUEpred),sd.obs,log=TRUE))
      #suma de -LogLike
      return(Lt)
    }

    r[j] = aux$r[j]
```

```

K[j] = aux$K[j]
if(!is.na(loglik2011(c(r[j],K[j])))){
  optimo = optim(c(r[j],K[j]), loglik2011, hessian=TRUE)
  #función que optimiza
  aux$r.optimo[j] = round(optimo$par[1],3)
  aux$K.optimo[j] = round(optimo$par[2])
  aux$sumLL[j] = round(optimo$value,4)
}

if(is.na(loglik2011(c(r[j],K[j])))){
  aux$r.optimo[j] = NA
  aux$K.optimo[j] = NA
  aux$sumLL[j] = NA
}
}
}

# Error en el proceso
if(!is.null(q.inicial)){
  q = seq(q.inicial, q.final, q.intervalo)
  aux = expand.grid(r = r, K = K, q = q)
  aux$r.optimo = rep(0, nrow(aux))
  aux$K.optimo = rep(0, nrow(aux))
  aux$q.optimo = rep(0, nrow(aux))
  aux$sumLL = rep(0, nrow(aux))
  for (j in 1:nrow(aux)){
    print(j)
    loglik2011 <- function(x){
      aux$r[j]<-x[1]#valor inicial de r
      aux$K[j]<-x[2]#valor inicial de K
      aux$q[j]<-x[3]#valor inicial de q

```

```

Bpred <- numeric(n)
#Biomasa predicha por el modelo de Schaefer
Bpred[1] = datos[1,2]/aux$q[j]
for (i in 2:n){
  Bpred[i] = (datos[i-1,2]/aux$q[j]) +
  (aux$r[j]*(datos[i-1,2]/aux$q[j]))*(1-(datos[i-1,2]/
  (aux$q[j]*aux$K[j]))) - datos[i-1,3] #modelo de Schaefer
}

CPUEpred = Bpred*aux$q[j] #CPUE predicha por el modelo
sd.proc = sqrt(sum(((log(datos[,2])-log(CPUEpred))^2)/(n-1)))
Lt = sum(-dnorm(log(datos[,2]), log(CPUEpred), sd.proc,
log = TRUE)) #suma de -LogLike
return(Lt)
}

r[j] = aux$r[j]
K[j] = aux$K[j]
q[j] = aux$q[j]
if(!is.na(loglik2011(c(r[j],K[j],q[j])))){
  optimo = optim(c(r[j],K[j],q[j]), loglik2011)
  #función que optimiza
  aux$r.optimo[j] = round(optimo$par[1],3)
  aux$K.optimo[j] = round(optimo$par[2])
  aux$q.optimo[j] = optimo$par[3]
  aux$sumLL[j] = round(optimo$value,4)
}

if(is.na(loglik2011(c(r[j],K[j],q[j])))){
  aux$r.optimo[j] = NA
  aux$K.optimo[j] = NA
  aux$q.optimo[j] = NA
  aux$sumLL[j] = NA
}

```

```

}
}
#La salida "optimo" permite observar información sobre la optimización,
#pero es de especial interés $par y $value que corresponden a los
#optimos de r, K, q y el valor de la verosimilitud maximizada.

if(is.null(Bo)){Tabla = aux
Tabla = as.data.frame(Tabla)
}
if(!is.null(Bo)){Tabla = cbind(Bo = rep(Bo, nrow(aux)), aux)
Tabla = as.data.frame(Tabla)
Tabla = Tabla[which(Tabla$K.optimo >= Bo),]
Tabla$RMS.optimo = round((Tabla$r.optimo*Tabla$K.optimo)/4)
Tabla = Tabla[which(Tabla$RMS.optimo >= 0),]
return(Tabla)
}

#####

resultados <- function(r, K, q.proc = NULL, Bo = NULL)
{
  optimos = subopt
  n = nrow(datos) #cantidad de años
  Bpred = c()
  if(is.null(q.proc)){
    if(is.null(Bo)){Bpred[1] = K}
    if(!is.null(Bo)){Bpred[1] = Bo}
    for (i in 2:n)
      {
        Bpred[i] = Bpred[i-1]+((r*Bpred[i-1])*(1-
          (Bpred[i-1]/K)))-datos[i-1,3]
      }
  }
}

```

```

    }
Log = log(datos[,2]/Bpred)
q = exp((1/n)*sum(Log))
Biobs = datos[,2]/q
CPUEpred = Bpred*q
logCPUEpred = log(CPUEpred)
sd.obs = sqrt(sum(((log(datos[,2]) -
log(CPUEpred))^2)/(n-1)))
if(!is.null(Bo)){
res = as.data.frame(cbind(optimos[which
(optimos$Bo == Bo & optimos$r.optimo == r &
optimos$K.optimo == K)], sd.obs, q, KS.test = ks.test(
(log(datos[,2]) - logCPUEpred),"pnorm", 0,
sd.obs)$p.value, SW.test = shapiro.test((log(datos[,2])
- logCPUEpred))$p.value))
aux = as.data.frame(cbind(Bo = rep(Bo, length(Bpred)),
r = rep(r, length(Bpred)), K = rep(K, length(Bpred)),
Año = datos[,1], CPUEobs = datos[,2],
CPUEpred = round(CPUEpred,3), Biobs = round(Biobs),
Bpred = round(Bpred)))
write.table(aux, paste('Ajuste', Bo, round(r,2),
round(K), '.csv', sep = ' '), sep = ';', dec = ',',
row.names = FALSE)
}
if(is.null(Bo)){
res = as.data.frame(cbind(optimos[which(optimos$r.optimo == r
& optimos$K.optimo == K)], sd.obs, q,
KS.test = ks.test((log(datos[,2]) - logCPUEpred),
"pnorm", 0, sd.obs)$p.value,
SW.test = shapiro.test((log(datos[,2]) - logCPUEpred))$p.value))
aux = as.data.frame(cbind(r = rep(r, length(Bpred)),

```

```

K = rep(K, length(Bpred)), Año = datos[,1],
CPUEobs = datos[,2], CPUEpred = round(CPUEpred,3),
  Biobs = round(Biobs), Bpred = round(Bpred))
write.table(aux, paste('Ajuste', round(r,2), round(K),
  '.csv', sep = ' '), sep = ';', dec = ',', row.names = FALSE)
}

#Gráfica CPUE predicha y CPUE observada
if(is.null(Bo)){pdf(paste('Ajustecpue', '.pdf', sep = ''))}
if(!is.null(Bo)){pdf(paste('AjustecpueBo', Bo, '.pdf', sep = ''))}
par(mfrow=c(1,2))
plot(datos[,1], datos[,2], "p", pch = 20, xlab = "Año",
ylab = "CPUE", ylim = c(0,max(datos[,2])))
lines(datos[,1], CPUEpred, col = "blue")
legend(col = c("black","blue"),"bottomleft",bty = "n",
  pch = c(19,20),c('CPUE observada','CPUE predicha'),
  lty=c(0,1), cex=0.5)

#Gráfica Biomasa predicha y Biomasa observada
plot(datos[,1], Biobs, "p", pch = 20, xlab = "Año",
ylab = "Biomasa", ylim = c(0,max(Biobs)))
lines(datos[,1], Bpred, col = "blue")
legend(col = c("black", "blue"), "bottomleft",
bty = "n", pch = c(19,20), c('Biomasa observada',
'Biomasa predicha'),lty=c(0,1), cex=0.5)
dev.off()
}

if(!is.null(q.proc)){
Bpred[1] = datos[1,2]/q.proc
for (i in 2:n)

```



```

{
  Bpred[i] = (datos[i-1,2]/q.proc)+((r*datos[i-1,2])/q.proc)*(1-
    (datos[i-1,2]/(q.proc*K))) - datos[i-1,3]
# Bpred[i] = Bpred[i]+((r*Bpred[i])/q)*(1-(datos[i-1,2]/K))
- datos[i-1,3]

}

Biobs = datos[,2]/q.proc
CPUEpred = Bpred*q.proc
logCPUEpred = log(CPUEpred)
sd.proc = sqrt(sum(((log(datos[,2])-log(CPUEpred))^2)/(n-1)))
res = as.data.frame(cbind(optimos[which(optimos$r.optimo == r &
  optimos$K.optimo == K & optimos$q.optimo == q.proc ),],
sd.proc, KS.test = ks.test((log(datos[,2]) - logCPUEpred),
  "pnorm", 0, sd.proc)$p.value,
  SW.test = shapiro.test((log(datos[,2]) - logCPUEpred))$p.value))
aux = as.data.frame(cbind(r = rep(r, length(Bpred)),
K = rep(K, length(Bpred)), Año = datos[,1],
CPUEobs = datos[,2], CPUEpred = CPUEpred,
Biobs = Biobs, Bpred = Bpred))
write.table(aux, paste('Ajuste proc', round(r,2),
round(K), q.proc, '.csv', sep = ' '), dec = ',',
sep = ';', row.names =FALSE)

pdf(paste('Ajustecpue_proc', '.pdf', sep = ''))
par(mfrow=c(1,2))
plot(datos[,1], datos[,2], "p", pch = 20, xlab = "Año",
ylab = "CPUE", ylim = c(0, max(datos[,2])))
lines(datos[,1], CPUEpred, col = "blue")
legend(col = c("black", "blue"), "bottomleft", bty = "n",
pch = c(19,20), c('CPUE observada', 'CPUE predicha'),

```

```

lty = c(0, 1), cex = 0.5)

#Gráfica Biomasa predicha y Biomasa observada
plot(datos[,1], Biobs, "p", pch = 20, xlab = "Año",
ylab = "Biomasa", ylim = c(0, max(Biobs)))
lines(datos[,1], Bpred, col = "blue")
legend(col = c("black", "blue"), "bottomleft",
bty = "n", pch = c(19, 20), c('Biomasa observada',
'Biomasa predicha'), lty = c(0, 1), cex = 0.5)
dev.off()
}

return(res)
}

#####

nivel = function(inf.K, sup.K, salto.K, inf.r, sup.r,
salto.r, Bo = NULL, q.proc = NULL)
{
K = seq(inf.K,sup.K,salto.K)
# Se evalúa la verosimilitud en este rango de K
r = seq(inf.r,sup.r,salto.r)
# Se evalúa la verosimilitud en este rango de r
n = nrow(datos)

if(is.null(q.proc)){
aux = expand.grid(r = r, K = K)
if(!is.null(Bo)){aux = aux[which(aux$K >= Bo),]}
# Solo toma para cada Bo los K>=Bo
aux$Sum_Log_Like = rep(0, nrow(aux))

```

```

for (j in 1:nrow(aux)){
  Bpred = c()
  if(!is.null(Bo)){Bpred[1] = Bo}
  if(is.null(Bo)){Bpred[1] = aux$K[j]}
for (i in 2:nrow(datos))
  {
    Bpred[i] = Bpred[i-1]+((aux$r[j]*Bpred[i-1])*
      (1-(Bpred[i-1]/aux$K[j]))) - datos[i-1,3]
  }
  Log <- (log(datos[,2]/Bpred))
  q <- exp((1/n)*sum(Log)) #Coeficiente de capturabilidad
  Biobs = datos[,2]/q
  CPUEpred = Bpred*q
  sd.obs = sqrt(sum(((log(datos[,2]) -
    log(CPUEpred))^2)/(n-1))) # desvío de las obs
  Lt = -dnorm(log(datos[,2]), log(CPUEpred),
    sd.obs, log = TRUE)
  aux$Sum_Log_Like[j] = sum(Lt)
}
}

if(!is.null(q.proc)){
q = q.proc # Se evalúa la verosimilitud en este rango de q
aux = expand.grid(r = r, K = K, q = q)
aux$Sum_Log_Like = rep(0, nrow(aux))
for (j in 1:nrow(aux)){
  Bpred = c()
  Bpred[1] = datos[1,2]/aux$q[j]
  for (i in 2:nrow(datos))
    {
      Bpred[i] = (datos[i-1,2]/aux$q[j]) +
        (aux$r[j]*(datos[i-1,2]/aux$q[j]))*(1-

```

```

    (datos[i-1,2]/(aux$q[j]*aux$K[j]))) - datos[i-1,3]
    #modelo de Schaefer
  }
  CPUEpred = Bpred*aux$q[j] #CPUE predicha por el modelo
  sd.proc = sqrt(sum(((log(datos[,2])-log(CPUEpred))^2)/(n-1)))
  Lt = -dnorm(log(datos[,2]), log(CPUEpred),
  sd.proc, log = TRUE)
  aux$Sum_Log_Like[j] = sum(Lt)
}
}

aux1 = na.omit(aux)
#Se re-escala la función para que esté entre 0 y 1
aux1$norm.sum.LL = (aux1$Sum_Log_Like -
max(aux1$Sum_Log_Like))/min(aux1$Sum_Log_Like -
max(aux1$Sum_Log_Like))

if(!is.null(Bo)){
  aux1$Bo = Bo
  pdf(paste('nivelBo', Bo, '.pdf', sep = ''))
  print(levelplot(norm.sum.LL ~ r*K, data = aux1,
xlab = "r", ylab = "K", cuts = 15,
main = paste('Bo = ', Bo, sep = ''),
col.regions = terrain.colors,
colorkey = T, region = TRUE))
dev.off()
}

if(is.null(Bo)){
  if(is.null(q.proc)){pdf(paste('nivel',
'.pdf', sep = ''))}
  if(!is.null(q.proc)){pdf(paste('nivel',

```

```

q.proc, '.pdf', sep = ''))}
print(levelplot(norm.sum.LL ~ r*K, data =
  aux1, xlab = "r", ylab = "K", cuts = 15,
col.regions = terrain.colors, colorkey = T,
  region = TRUE))
dev.off()
}
return(aux1)
}

#####

perfilr = function(inf.r, sup.r, salto.r,
K, Bo = NULL, q.proc = NULL)
{
n = nrow(datos)
r.seq = seq(inf.r, sup.r, salto.r)
# Se evalúa la verosimilitud en este rango de r
vero.perfil.r = c()

if(is.null(q.proc)){
for(t in 1:length(r.seq)){
loglik2011.K = function(K){
Bpred = numeric(n)
#Biomasa predicha por el modelo de Schaefer
if(!is.null(Bo)){Bpred[1] = Bo}
if(is.null(Bo)){Bpred[1] = K}
for (i in 2:n){
Bpred[i] = Bpred[i-1]+((r.seq[t]*Bpred[i-1])*(1-
(Bpred[i-1]/K)))-datos[i-1,3] #modelo de Schaefer
}

```

```

Log <-(log(datos[,2]/Bpred))
q <- exp((1/n)*sum(Log)) #Coeficiente de capturabilidad
CPUEpred = Bpred*q #CPUE predicha por el modelo
sd.obs = sqrt(sum(((log(datos[,2])-
log(CPUEpred))^2)/(n-1)))
Lt = sum(-dnorm(log(datos[,2]),log(CPUEpred),
sd.obs,log=TRUE)) #suma de -LogLike
return(Lt)
}
if(!is.na(loglik2011.K(K))){
optimo.K = optim(K, loglik2011.K, hessian = TRUE)
#función que optimiza
vero.perfil.r[t] = optimo.K$value
#La verosimilitud perfil de r
}
if(is.na(loglik2011.K(K))){vero.perfil.r[t] = NA}
}
}

if(!is.null(q.proc))
{
for(t in 1:length(r.seq)){
print(t)
loglik2011.K = function(x){
K = x[1]
q.proc = x[2]
Bpred = numeric(n)
#Biomasa predicha por el modelo de Schaefer
Bpred[1] = datos[1,2]/q.proc
for (i in 2:n){
Bpred[i] = (datos[i-1,2]/q.proc)+(r.seq[t]*(datos[i-1,2]

```

```

/q.proc))*(1-(datos[i-1,2]/(q.proc*K))) -
datos[i-1,3] #modelo de Schaefer
}
CPUEpred = Bpred*q.proc #CPUE predicha por el modelo
sd.proc = sqrt(sum(((log(datos[,2]) -
log(CPUEpred))^2)/(n-1)))
Lt = sum(-dnorm(log(datos[,2]), log(CPUEpred),
sd.proc, log = TRUE)) #suma de -LogLike
return(Lt)
}

if(!is.na(loglik2011.K(c(K, q.proc)))){
optimo.K = optim(c(K,q.proc), loglik2011.K)
#función que optimiza
vero.perfil.r[t] = optimo.K$value
#La verosimilitud perfil de r
}
if(is.na(loglik2011.K(c(K,q.proc)))){vero.perfil.r[t] = NA}
}
}

r.perf = as.data.frame(cbind(r.seq, vero.perfil.r))
# Sacamos los NA's
r.perf = na.omit(r.perf)

#Distancia chi cuadrado para obtener
#el intervalo de confianza de r
r.perf$dist.chi.r = 2*(r.perf[,2] - min(r.perf[,2]))
r.perf$pchi.r = pchisq(r.perf$dist.chi.r,1)
r.perf$dchir = dchisq(r.perf$dist.chi.r,1)

```

```
#r.chiq = cbind(r.perf,chi.r)
#int.conf.r = which(round(r.chiq[,3],2)<(1-alfa))
#int.conf.r = r.chiq[int.conf.r,]
#inf.interval.conf.r =
int.conf.r[which(int.conf.r[,3] == min(int.conf.r[,3])),1]
#sup.interval.conf.r =
int.conf.r[which(int.conf.r[,3] == max(int.conf.r[,3])),1]
#inf.interval.conf.r
#sup.interval.conf.r

#Intervalo.de.confianza.de.r =
cbind(inf.interval.conf.r,sup.interval.conf.r, Bo)

if(is.null(q.proc))
{
if(!is.null(Bo)){
r.perf$Bo = Bo
pdf(paste('perfilrBo', Bo, '.pdf', sep = ''))
}
if(is.null(Bo)){pdf(paste('perfilr', '.pdf', sep = ''))}
par(mfrow = c(1,2))
plot(r.perf[,1], r.perf[,2], "l",
ylab = expression(-LL(r)), main = 'a'),
xlab = "r",frame = FALSE, cex = 1.5)
plot(r.perf[,1], r.perf[,4], "l", xlab = "r",
ylab = expression(paste('Probabilidad ', chi^2, sep = ' ')),
main = 'b'),lwd = 2, cex = 1.5)
dev.off()
}

if(!is.null(q.proc))
```



```
{
pdf(paste('perfilr_proc', '.pdf', sep = ''))
par(mfrow = c(1,2))
plot(r.perf[,1], r.perf[,2], "l",
ylab = expression(-LL(r)),main = 'a'),
xlab = "r",frame = FALSE, cex = 1.5)
plot(r.perf[,1], r.perf[,4], "l", xlab = "r",
ylab = expression(paste('Probabilidad ', chi^2, sep = ' ')),
main = 'b'),lwd = 2, cex = 1.5)
dev.off()
}

return(r.perf)
}
```

```
#####
```

```
perfilK = function(inf.K, sup.K, salto.K,
r, Bo = NULL, q.proc = NULL)
{
n = nrow(datos)
K.seq = seq(inf.K, sup.K, salto.K)
# Se evalúa la verosimilitud en este rango de r
vero.perfil.K = c()

if(is.null(q.proc))
{
for(t in 1:length(K.seq)){
loglik2011.r = function(r){
Bpred = numeric(n)
#Biomasa predicha por el modelo de Schaefer
```

```

    if(!is.null(Bo)){Bpred[1] = Bo}
  if(is.null(Bo)){Bpred[1] = K.seq[t]}
  for (i in 2:n){
    Bpred[i] = Bpred[i-1]+((r*Bpred[i-1])*
      (1-(Bpred[i-1]/K.seq[t])))-datos[i-1,3]
    #modelo de Schaefer
  }
  Log <- (log(datos[,2]/Bpred))
  q <- exp((1/n)*sum(Log)) #Coeficiente de capturabilidad
  CPUEpred = Bpred*q #CPUE predicha por el modelo
  sd.obs = sqrt(sum(((log(datos[,2])-log(CPUEpred))^2)/(n-1)))
  Lt = sum(-dnorm(log(datos[,2]),
    log(CPUEpred),sd.obs,log=TRUE)) #suma de -LogLike
  return(Lt)
}

if(!is.na(loglik2011.r(r))){
  optimo.r = optim(r, loglik2011.r, hessian = TRUE)
  #función que optimiza
  vero.perfil.K[t] = optimo.r$value
  #La verosimilitud perfil de r
}
if(is.na(loglik2011.r(r))){vero.perfil.K[t] = NA}
}
}

if(!is.null(q.proc))
{
  for(t in 1:length(K.seq)){
    print(t)
    loglik2011.r = function(x){
      r = x[1]

```

```

q.proc = x[2]
  Bpred = numeric(n)
  #Biomasa predicha por el modelo de Schaefer
  Bpred[1] = datos[1,2]/q.proc
  for (i in 2:n){
Bpred[i] = (datos[i-1,2]/q.proc)+(r*(datos[i-1,2]/
q.proc))*(1-(datos[i-1,2]/(q.proc*K.seq[t])))
- datos[i-1,3] #modelo de Schaefer
  }
  CPUEpred = Bpred*q.proc
  #CPUE predicha por el modelo
  sd.proc = sqrt(sum(((log(datos[,2])-
log(CPUEpred))^2)/(n-1)))
  Lt = sum(-dnorm(log(datos[,2]),log(CPUEpred),
sd.proc,log=TRUE)) #suma de -LogLike
  return(Lt)
}

if(!is.na(loglik2011.r(c(r, q.proc)))){
optimo.r = optim(c(r, q.proc), loglik2011.r)
#función que optimiza
vero.perfil.K[t] = optimo.r$value
#La verosimilitud perfil de r
}
if(is.na(loglik2011.r(c(r,q.proc)))){vero.perfil.K[t] = NA}
}
}

perf.K = as.data.frame(cbind(K.seq, vero.perfil.K))
# Sacamos los NA's
perf.K = na.omit(perf.K)

```

```
#Distancia chi cuadrado para obtener
#el intervalo de confianza de r
perf.K$dist.chi.K = 2*(perf.K[,2] - min(perf.K[,2]))
perf.K$pchi.K = pchisq(perf.K$dist.chi.K,1)
perf.K$dchi.K = dchisq(perf.K$dist.chi.K,1)

#K.chiq = cbind(vero.perfil.K,chi.K)
#int.conf.K = which(round(K.chiq[,2],2)<(1-alfa))
#int.conf.K = K.chiq[int.conf.K,]
#inf.interval.conf.K =
int.conf.K[which(int.conf.K[,3] == min(int.conf.K[,3])),1]
#sup.interval.conf.K =
int.conf.K[which(int.conf.K[,3] == max(int.conf.K[,3])),1]
#inf.interval.conf.K
#sup.interval.conf.K

#Intervalo.de.confianza.de.K =
  cbind(inf.interval.conf.K, sup.interval.conf.K, Bo)

if(is.null(q.proc))
{
  if(!is.null(Bo)){
    perf.K$Bo = Bo
    pdf(paste('perfilKBo', Bo, '.pdf', sep = ''))
  }
  if(is.null(Bo)){pdf(paste('perfilK', '.pdf', sep = ''))}
  par(mfrow = c(1,2))
  plot(perf.K[,1], perf.K[,2], "l",
  ylab = expression(-LL(K)), xlab = "K", frame = FALSE)
  plot(perf.K[,1], perf.K[,4], "l", xlab = "K",
  ylab = expression(paste('Probabilidad ',
```

```

chi^2, sep = ' '),lwd = 2)
dev.off()
}

if(!is.null(q.proc))
{
pdf(paste('perfilK_proc', '.pdf', sep = ''))
par(mfrow = c(1,2))
plot(perf.K[,1], perf.K[,2], "l", ylab =
expression(-LL(K)), xlab = "K", frame = FALSE)
plot(perf.K[,1], perf.K[,4], "l", xlab = "K",
ylab = expression(paste('Probabilidad ', chi^2, sep = ' ')),
lwd = 2)
dev.off()
}
return(perf.K)
}

#####

perfilRMS = function(inf.RMS, sup.RMS, salto.RMS,
r, Bo = NULL, q.proc = NULL)
{
n = nrow(datos)
RMS.seq = seq(inf.RMS, sup.RMS, salto.RMS)
# Se evalúa la verosimilitud en este rango de r
vero.perfil.RMS = c()

if(is.null(q.proc))
{
for(t in 1:length(RMS.seq)){

```

```

print(t)
loglik2011.RMS = function(r){
  Bpred = numeric(n)
  #Biomasa predicha por el modelo de Schaefer
  if(!is.null(Bo)){Bpred[1] = Bo}
  if(is.null(Bo)){Bpred[1] = (4*RMS.seq[t])/r}
  for (i in 2:n){
    Bpred[i] = Bpred[i-1]+((r*Bpred[i-1])*
      (1-(Bpred[i-1]/((4*RMS.seq[t])/r))))-datos[i-1,3]
    #modelo de Schaefer
  }
  Log <- (log(datos[,2]/Bpred))
  q <- exp((1/n)*sum(Log)) #Coeficiente de capturabilidad
  CPUEpred = Bpred*q #CPUE predicha por el modelo
  sd.obs = sqrt(sum(((log(datos[,2])-
    log(CPUEpred))^2)/(n-1)))
  Lt = sum(-dnorm(log(datos[,2]), log(CPUEpred),
    sd.obs, log = TRUE)) #suma de -LogLike
  return(Lt)
}
if(!is.na(loglik2011.RMS(r))){
  optimo.RMS = optim(r, loglik2011.RMS, hessian = TRUE)
  #función que optimiza
  vero.perfil.RMS[t] = optimo.RMS$value
  #La verosimilitud perfil de r
}
if(is.na(loglik2011.RMS(r))){vero.perfil.RMS[t] = NA}
}
}

if(!is.null(q.proc))

```

```
{
for(t in 1:length(RMS.seq)){
print(t)
loglik2011.RMS = function(x){
r = x[1]
q.proc = x[2]
  Bpred = numeric(n) #Biomasa predicha por el modelo de Schaefer
  Bpred[1] = datos[,2]/q.proc
  for (i in 2:n){
    Bpred[i] = (datos[i-1,2]/q.proc)+(r*(datos[i-1,2]/
      q.proc))*(1-r*(datos[i-1,2]/((q.proc*RMS.seq[t])*4)))
    - datos[i-1,3] #modelo de Schaefer
  }
  CPUEpred = Bpred*q.proc #CPUE predicha por el modelo
  sd.proc = sqrt(sum(((log(datos[,2])-
    log(CPUEpred))^2)/(n-1)))
  Lt = sum(-dnorm(log(datos[,2]), log(CPUEpred),
    sd.proc, log = TRUE)) #suma de -LogLike
  return(Lt)
}

if(!is.na(loglik2011.RMS(c(r, q.proc)))){
optimo.RMS = optim(c(r, q.proc), loglik2011.RMS)
#función que optimiza
vero.perfil.RMS[t] = optimo.RMS$value
#La verosimilitud perfil de r
}
if(is.na(loglik2011.RMS(c(r, q.proc)))){
vero.perfil.RMS[t] = NA}
}
}
```

```
perf.RMS = as.data.frame(cbind(RMS.seq, vero.perfil.RMS))
perf.RMS = na.omit(perf.RMS)
```

```
#Distancia chi cuadrado para obtener
#el intervalo de confianza de r
perf.RMS$dist.chi.RMS = 2*(perf.RMS[,2]-
min(perf.RMS[,2]))
perf.RMS$pchi.RMS = pchisq(perf.RMS$dist.chi.RMS,1)
perf.RMS$dchi.RMS = dchisq(perf.RMS$dist.chi.RMS,1)
```

```
if(is.null(q.proc))
{
  if(!is.null(Bo)){
    perf.RMS$Bo = Bo
    pdf(paste('perfilRMSBo', Bo, '.pdf', sep = ''))
  }
  if(is.null(Bo)){pdf(paste('perfilRMS',
'.pdf', sep = ''))}
  par(mfrow=c(1,2))
  plot(perf.RMS[,1], perf.RMS[,2], "l",
ylab = expression(-LL(RMS)),xlab = "RMS", frame = FALSE)
  plot(perf.RMS[,1], perf.RMS[,4], "l",
  xlab = "RMS", ylab = expression(paste('Probabilidad ',
chi^2, sep = ' ')), lwd = 2)
  dev.off()
}
```

```
if(!is.null(q.proc))
{
  pdf(paste('perfilRMS_proc', '.pdf', sep = ''))
  par(mfrow=c(1,2))
```



```

plot(perf.RMS[,1], perf.RMS[,2], "l",
ylab = expression(-LL(RMS)), xlab = "RMS", frame = FALSE)
plot(perf.RMS[,1], perf.RMS[,4], "l",
xlab = "RMS", ylab = expression(paste('Probabilidad ',
chi^2, sep = ' ')), lwd=2)
dev.off()
}

#intervalo de confianza al 1-alfa para RMS
#int.conf.RMS = which(round(RMS.chiq[,3],2)<(1-alfa))
#int.conf.RMS = RMS.chiq[int.conf.RMS,]
#inf.interval.conf.RMS =
# int.conf.RMS[which(int.conf.RMS[,3] ==
# min(int.conf.RMS[,3])),1]
#sup.interval.conf.RMS =
#int.conf.RMS[which(int.conf.RMS[,3] ==
# max(int.conf.RMS[,3])),1]
#inf.interval.conf.RMS
#sup.interval.conf.RMS

return(perf.RMS)
}

#####

```

8.1.2. Datos de la literatura

```

#Datos
# Merluza de Namibia, Hilborn Pág. 240)

datos=list(Año=c(1965, 1966, 1967, 1968, 1969, 1970,

```

```
1971, 1972, 1973, 1974, 1975, 1976, 1977, 1978, 1979,
  1980, 1981, 1982, 1983, 1984, 1985, 1986, 1987),
CPUE = c(1.78, 1.31, 0.91, 0.96, 0.88, 0.90, 0.87,
0.72, 0.57, 0.45, 0.42, 0.42, 0.49, 0.43,
0.40, 0.45, 0.55, 0.53, 0.58, 0.64, 0.66, 0.65,0.63),
Captura = c(94, 212, 195, 383, 320, 402,
366, 606, 378, 319, 309, 389, 277, 254,
170, 97, 91, 177, 216, 229, 211, 231, 223))
datos=data.frame(datos)
```

8.1.3. Script principal para ajuste con incertidumbre en la observaciones

```
library(latticeExtra)
source('rangoptim.R')
source('resultados.R')
source('nivel.R')
source('perfilr.R')
source('perfilK.R')
source('perfilRMS.R')

#####
#####Cargar datos#####
#####

##Los datos deben estar ordenados de la siguiente manera
#AÑO, CPUE, CAPTURA
```

```
#####  
##Valores iniciales del modelo#####  
#####  
  
## Capacidad de carga "K"  
# Rango de valores iniciales para K  
  
K.inicial = 1000 # limite inferior para K  
K.final = 10000 # limite superior para K  
K.intervalo = 1000 # salto de K  
  
# Biomasa inicial "Bo"  
# B0 <= K  
# Para el caso de Hilborn Bo = K, entonces Bo = NULL  
# porque la biomasa inicial es K  
  
## Tasa intrínseca de crecimiento ?r?  
# Rango de valores iniciales para r  
  
r.inicial = 0.1 # limite inferior para r  
r.final = 1 # limite superior para r  
r.intervalo = 0.1 # salto de r  
  
# Bajo el supuesto de incertidumbre en las observaciones  
# q analítico, entonces no se necesita valor  
# inicial para q, por defecto es NULL  
  
# La función "rangoptim" retorna una tabla con  
# los valores iniciales de Bo, en este caso igual a K,  
# r y K y los valores optimos de r y K
```

```
# a partir de dichos valores iniciales

optimos = rangoptim(K.inicial, K.final, K.intervalo,
r.inicial, r.final, r.intervalo)

summary(optimos) # vemos la variabilidad de los resultados

# Para cualquier combinación de r y K inicial
#los valores óptimos de estos
# no varían

#####
####Gráfica de la verosimilitud####
#Superficies de nivel para r, K y L##
#####Variando r y K#####
#####

inf.K = 2000 #límite inferior para K
sup.K = 4000 #límite superior para K
salto.K = 10 #salto de K

inf.r = 0.2 #límite inferior para r
sup.r = 0.6 #límite superior para r
salto.r = 0.01 #salto de r

print(nivel(inf.K, sup.K, salto.K, inf.r, sup.r, salto.r))

#####
####Resultados de interés#####
#####
```

```
# Con los óptimos obtenidos de r y K
# La función resultados da los valores de interés
# Es conveniente optar por alguno de los posibles
# optimos que proporciona la función, porque para
#cada uno guarda una gráfica y sino se selecciona
#un subconjunto son muchos

subopt = optimos[which(optimos$r == 1 & optimos$K == 10000),]
res = resultados(subopt$r.optimo, subopt$K.optimo)
res

#####
#####Perfiles de r #####
#####

alfa = 0.05
#Nivel de significación para el intervalos de confianza de r

inf.r = 0.1 #límite inferior para r
sup.r = 0.9  #límite superior para r
salto.r = 0.001  #salto de r

K = 10000

perf.r = perfilr(inf.r, sup.r, salto.r, K)

#####
#####Perfiles de K #####
#####
```

```

alfa = 0.05
#Nivel de significación para el intervalos de confianza de K

inf.K = 1000
sup.K = 6000 #límite superior para K
salto.K = 13 #salto de K
r = 2 #Valor inicial de r

perf.K = perfilK(inf.K, sup.K, salto.K, r)

#####
#####Perfiles de RMS #####
#####

alfa = 0.05
#Nivel de significación para el intervalos de confianza de K

inf.RMS = 240 #límite inferior para RMS
sup.RMS = 300 #límite superior para RMS
salto.RMS = 11 #salto de RMS
RMS.seq = seq(inf.RMS, sup.RMS, salto.RMS)
# Se evalúa la verosimilitud en este rango de r
r = .1
#Valor inicial de r, tiene que ser pequeño porque
# el termino #(1-(Bpred[i-1]/((4*RMS.seq[t])/r))))
#queda negativo y dan muchos na's

#Usamos la misma función del perfil de K y
#sustituimos K=(4*RMS.seq)/r

perf.RMS = perfilRMS(inf.RMS, sup.RMS, salto.RMS, r)

```

perf.RMS

8.1.4. Script principal para ajuste con incertidumbre en el proceso

```
library(latticeExtra)
source('rangoptim.R')
source('resultados.R')
source('nivel.R')
source('perfilr.R')
source('perfilK.R')
source('perfilRMS.R')
```

```
#####
#####Cargar datos#####
#####
```

```
##Los datos deben estar ordenados de la siguiente manera
#AÑO, CPUE (ton/hs), CAPTURA (ton)
```

```
#####
##Valores iniciales del modelo#####
#####
```

```
#1) Capacidad de carga "K"
# K= Valor inicial de K
K.inicial = 1000
```

```
K.final = 10000
K.intervalo = 1000

#2)Biomasa inicial "Bo"
# B0 < K
# parámetro opcional

# Bo = NULL

#3) Tasa intrínseca de crecimiento "r"

#r = Valor inicial de r

r.inicial = 0.1
r.final = 1
r.intervalo = 0.1

# Valor inicial de q
q.inicial = 0.0001
q.final = 0.00045 # es el q que da con error en las obs
q.intervalo = 0.0001

#4) Valor inicial del desvío de las observaciones "sd.obs"
# estos parámetros son opcionales si no
#se corren entonces por defecto son nulos

#sd.proc = 1 # porque es con erro en las observaciones

#La función "rangoptim" da una tabla con los valores
# iniciales de Bo, r y K y los distintos optimos de r y K
```



```
optimos = rangoptim(K.inicial, K.final, K.intervalo,
  r.inicial, r.final, r.intervalo, q.inicial = q.inicial,
  q.final = q.final, q.intervalo = q.intervalo)
```

```
summary(optimos)
```

```
# Vemos que para cualquier combinación
#de r y K optimiza en la misma solución
```

```
#####
####Gráfica de la verosimilitud#####
#Superficies de nivel para r, K y L##
#####Variando r y K#####
#####
```

```
inf.K = 2000 #límite inferior para K
sup.K = 4000 #límite superior para K
salto.K = 10 #salto de K
```

```
inf.r = 0.1 #límite inferior para r
sup.r = 0.6 #límite superior para r
salto.r = 0.01 #salto de r
```

```
q.proc = c(min(optimos$q.optimo),
  median(optimos$q.optimo), max(optimos$q.optimo))
explora = list()
for(i in 1:length(q.proc)){
  explora[[i]] = nivel(inf.K, sup.K, salto.K,
  inf.r, sup.r, salto.r, q.proc = q.proc[i])
```

```

}
explora1 = do.call(rbin, explora)
summary(explora)
#####
#####Resultados de interés#####
#####

# Con los óptimos obtenidos de r y K
# La función resultados da los valores de interés
# Es conveniente optar por alguno de los posibles
#óptimos que proporciona la función,
#porque para cada uno guarda una gráfica y
#sino se selecciona un subconjunto son muchos

subopt = optimos[which.min(optimos$sumLL),]
res = resultados(subopt$r.optimo, subopt$K.optimo, q = subopt$q.optimo)


#####
#####Perfiles de r #####
#####

alfa = 0.05
#Nivel de significación para el intervalos de confianza de r

inf.r = 0.1 #límite inferior para r
sup.r = .9  #límite superior para r
salto.r = 0.001  #salto de r

K = 1000000

```

```
# grande para evitar problemas numéricos
#porque q es pequeño
q = res$q.optimo
# para evitar problemas numéricos

perf.r = perfilr(inf.r,sup.r,salto.r,K, q.proc = q)

#####
#####Perfiles de K #####
#####

alfa=0.05
#Nivel de significación para el intervalos de confianza de K

inf.K = 500
sup.K = 6000 #límite superior para K
salto.K = 11 #salto de K
r = res$r.optimo #Valor inicial de r
q = res$q.optimo # para evitar problemas numéricos

q = .1 # para evitar problemas numéricos

perf.K = perfilK(inf.K, sup.K, salto.K, r, q.proc = q)

#####
#####Perfiles de RMS #####
#####

alfa = 0.05
#Nivel de significación para el intervalos de confianza de K
```

```

inf.RMS = 50 #limite inferior para RMS
sup.RMS = 1000 #límite superior para RMS
salto.RMS = 1 #salto de RMS
#RMS.seq = seq(inf.RMS, sup.RMS, salto.RMS)
# Se evalúa la verosimilitud en este rango de r
r = .1
#Valor inicial de r, tiene que ser pequeño porque
#sino el termino (1-(Bpred[i-1]/((4*RMS.seq[t])/r))))
# queda negativo y dan muchos na's

#Usamos la misma función del perfil de K
#y sustituimos K=(4*RMS.seq)/r

perf.RMS = perfilRMS(inf.RMS, sup.RMS, salto.RMS, r, q.proc = q)
perf.RMS

```

8.2. Script de R para estimar parámetros de distribuciones a priori

Ejemplo de script para al ajuste con datos de la literatura.

```

#Script para encontrar parámetros de la a priori de r, K
rm(list=ls(all=TRUE))

#####
##### r #####
#####
#Intervalo de r

```

```

rinf=0.1
rsup=0.8
#alfa.r es la probabilidad que queda afuera del intervalo
#definido por rinf y rsup
alfa.r=0.1
#Probabilidad acumulada de Kinf
zinf.r=qnorm(alfa.r)
#Probabilidad acumulada de Ksup
zsup.r=qnorm(1-alfa.r)
#Parámetros de la lognormal
sigma.r=(log(rsup)-log(rinf))/(zsup.r-zinf.r)
mu.r=(log(rinf)-zinf.r*sigma.r)
prec.r=1/(sigma.r^2)
parametros.r=c(mu.r,sigma.r,prec.r,alfa.r,rinf,rsup)
names(parametros.r)=c("Mu","Sigma","Precisión","alfa","rinf","rsup")
parametros.r

#####
#####      K #####
#####

#Intervalo de K
Kinf=1000
Ksup=8000
#alfaK es la probabilidad que queda afuera del intervalo
#definido por Kinf y Ksup
alfa.K=0.1
#Probabilidad acumulada de Kinf
zinf.K=qnorm(alfa.K)
#Probabilidad acumulada de Ksup
zsup.K=qnorm(1-alfa.K)

```

```
#Parámetros de la lognormal
sigma.K=(log(Ksup)-log(Kinf))/(zsup.K-zinf.K)
mu.K=(log(Kinf)-zinf.K*sigma.K)
prec.K=1/(sigma.K^2)
parametros.K=c(mu.K,sigma.K,prec.K,alfa.K,Kinf,Ksup)
names(parametros.K)=c("Mu","Sigma","Precisión","alfa","Kinf","Ksup")
parametros.K
```

8.3. Script de OpenBUGS para ajustar el modelo con el enfoque Bayesiano

Ejemplo de script para al ajuste con datos de la literatura.

```
#####
#1 Modelo
model:
{
# Distribución a priori de r: lognormal con 10% y 90% en 0.1 y 0.8
r ~ dlnorm(-1.2628643,1.5192832)I(0.1,0.8);

# Distribución a priori de K: lognormal con 10% y 90% en 1000 y 8000
K ~ dlnorm(7.9474760, 0.8112984)I(1000,8000);

# Distribución a priori de q: impropia (Prop. 1/q)
# utilizar sólo un adecuado intervalo para la inversa de "q" que es "iq"
iq ~ dgamma(0.001,0.001)I(1000,10000);
q <- 1/iq;
```

```
# Distribución a priori de sigma2: inv. gamma con 10% y 90% en 0.05 y 0.15
isigma2 ~ dgamma(1.708603,0.008613854);
sigma2 <- 1/isigma2;

# Distribución a priori de tau2: inv. gamma con 10% y 90% en 0.05 y 0.15
itau2 ~ dgamma(1.708603,0.008613854);
tau2 <- 1/itau2;

# (Condicional) Distribución a priori de los estados Ps:
Pmed[1] <- 0;
P[1] ~ dlnorm(Pmed[1],isigma2) I(0.001,2.0)
for (t in 2:N)
{
Pmed[t] <- log(max((P[t-1] + r*P[t-1]*(1-P[t-1]))- C[t-1]/K), 0.001));
P[t] ~ dlnorm(Pmed[t],isigma2)I(0.001,2.0)
}
# Distribución muestral:
for (t in 1:N)
{
CPUEmed[t] <- log(q*K*P[t]);
CPUE[t] ~ dlnorm(CPUEmed[t],itau2);
CPUEest[t] <- exp(CPUEmed[t])
}
# Otros parámetros de gestión y las predicciones:
RMS <- r*K/4;
}

#####

#2 Datos

list(N=23, C = c(94, 212, 195, 383, 320, 402, 366, 606, 378, 319, 309, 389,
277, 254,170, 97, 91, 177, 216, 229, 211, 231, 223), CPUE = c(1.78, 1.31,
```

```
0.91, 0.96, 0.88, 0.90, 0.87, 0.72, 0.57, 0.45, 0.42, 0.42, 0.49, 0.43,
0.40, 0.45, 0.55, 0.53, 0.58, 0.64, 0.66, 0.65,0.63))
```

```
#####
```

```
#3 Condiciones iniciales
```

```
# 1
```

```
list(P=c(0.99, 0.98, 0.96, 0.94, 0.92, 0.90, 0.88, 0.86, 0.84, 0.82,
0.80, 0.78, 0.76, 0.74, 0.72, 0.70, 0.68, 0.66, 0.64, 0.62, 0.60,
0.58, 0.56), r=0.5, K=3000, iq=5000, isigma2=100, itau2=100)
```

```
# 2
```

```
list(P=c(0.99, 0.98, 0.96, 0.94, 0.92, 0.90, 0.88, 0.86, 0.84, 0.82,
0.80, 0.78, 0.76, 0.74, 0.72, 0.70, 0.68, 0.66, 0.64, 0.62, 0.60,
0.58, 0.56), r=0.2, K=4000, iq=6000, isigma2=100, itau2=100)
```

```
# 3
```

```
list(P=c(0.99, 0.98, 0.96, 0.94, 0.92, 0.90, 0.88, 0.86, 0.84, 0.82,
0.80, 0.78, 0.76, 0.74, 0.72, 0.70, 0.68, 0.66, 0.64, 0.62, 0.60,
0.58, 0.56), r=0.3, K=5000, iq=2000, isigma2=100, itau2=100)
```

```
#####
```

```
#4 Correr
```

```
modelDisplay('log')
modelCheck ('C:/model.txt')
modelData('C:/datos.txt')
modelCompile(3)
modelInits('C:/condin1.txt')
modelInits('C:/condin2.txt')
modelInits('C:/condin3.txt')
```



```
samplesUpdate(100000)
```

```
samplesSet(r)
```

```
samplesSet(K)
```

```
samplesSet(RMS)
```

```
samplesSet(q)
```

```
samplesSet(CPUEest)
```

```
modelUpdate(300000)
```

```
samplesBeg(200000)
```

```
samplesStats('*')
```

```
samplesHistory('*')
```

```
samplesDensity('*')
```

```
samplesAutoC('*')
```

```
samplesQuantiles('*')
```