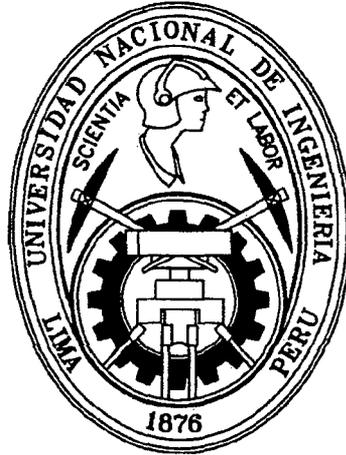


UNIVERSIDAD NACIONAL DE INGENIERÍA  
FACULTAD DE CIENCIAS  
ESCUELA PROFESIONAL DE MATEMÁTICA



Tesis para optar el título profesional de:  
LICENCIADO EN MATEMÁTICA

Titulada:

COMPARACIÓN DE LA EFICIENCIA DE  
DIFERENTES ALGORITMOS DE OPTIMIZACIÓN  
PARA LA SOLUCIÓN DE PROBLEMAS INVERSOS  
EN MODELOS DE DINÁMICA DE POBLACIONES

Presentado por:

Bach. Enrique Manuel Ramos Vásquez

Profesor Asesor:

Dr. Eladio Ocaña Anaya

Lima - Perú

2014

**Digitalizado por:**

Consortio Digital del  
Conocimiento MebLatam,  
Hemisferio y Dalse

**Dedicatoria**

**A mis padres y hermana,  
las personas que son la razón  
por la que puedo seguir adelante.**

## AGRADECIMIENTOS

Quiero agradecer al *Instituto del Mar del Perú (IMARPE)*, que me dio la oportunidad de hacer este trabajo y conocer temas nuevos en los cuales puedo aplicar las matemáticas.

A la *Dirección General de Investigaciones de Recursos Pelágicos*, en especial al *Área Funcional de Dinámica Poblacional y Evaluación de Recursos Pelágicos* y a todos los que la conforman, por su apoyo y confianza.

A los integrantes del *Centro de Investigaciones en Modelado Oceanográfico y Biológico Pesquero (CIMOBP)*, que me brindaron su amistad desde el inicio de mi etapa de tesista y me hicieron sentir como parte de ellos. Al Dr. Jorge Tam por sus palabras y consejos, tanto en el plano académico, como laboral y personal.

Al *Institut de Recherche pour le Développement (IRD)*, por haber financiado este trabajo. A Arnaud y Sophie, por su paciencia.

A Erich Díaz y Ricardo Oliveros, por su paciencia y dedicación, por su guía en la realización de esta tesis y los momentos de discusión compartidos.

Al Dr. Eladio Ocaña, por asesorarme en el desarrollo de este trabajo.

Y a todos aquellos que me dieron la oportunidad de conocer nuevos lugares donde aplicar las matemáticas.

## RESUMEN

En el presente trabajo se hace una revisión del proceso de formulación de modelos matemáticos empleados para la estimación de parámetros utilizados en dinámica de poblaciones, haciendo un especial énfasis en los métodos generalmente utilizados en la resolución de los problemas inversos en modelos de dinámica de poblaciones de peces.

En la primera parte del trabajo, se hace una clasificación de los modelos matemáticos (directo e inverso), centrándose el estudio en la solución de problemas inversos.

Se tratan los algoritmos de optimización tales como los métodos de *Newton*, *Quasi-Newton*, *Gradiente Conjugado*, además de algoritmos de optimización libres de derivadas.

Por otro lado, se estudia el proceso que muchas veces resulta no tan evidente, como es el proceso y la elección del método de estimación. Para ello, son considerados en este estudio los métodos de *mínimos cuadrados* y *máxima verosimilitud*.

Se hace una comparación de los métodos de optimización y estimación para determinar la mejor elección al momento de estimar los parámetros de un determinado modelo.

Finalmente, estudiamos una aplicación de lo anteriormente descrito, notando cómo incluso la elección de los datos influye en los resultados de forma significativa por lo que es importante considerarlo en el estudio.

# Índice general

<b>1. Introducción</b>	<b>1</b>
<b>2. Objetivos</b>	<b>3</b>
2.1. Objetivo general . . . . .	3
2.2. Objetivos específicos . . . . .	3
<b>3. Revisión de Literatura</b>	<b>4</b>
3.1. Modelo matemático . . . . .	4
3.1.1. Problema Directo . . . . .	4
3.1.2. Problema Inverso . . . . .	5
3.2. Fundamentos de optimización sin restricciones . . . . .	6
3.3. Métodos de optimización . . . . .	8
3.3.1. Métodos basados en derivadas . . . . .	8
3.3.1.1. Método de Newton . . . . .	11
3.3.1.2. Método Quasi-Newton . . . . .	12
3.3.1.3. Método de Gradiente Conjugado . . . . .	15
3.3.2. Métodos libres de derivadas . . . . .	17
3.3.2.1. Métodos deterministas . . . . .	18
3.3.2.2. Métodos estocásticos . . . . .	19
3.4. Métodos de estimación . . . . .	22
3.4.1. Método de mínimos cuadrados . . . . .	22
3.4.2. Método de máxima verosimilitud . . . . .	24
3.5. Método de Montecarlo . . . . .	24
3.6. Algunos modelos de dinámica de poblaciones de peces . . . . .	25
3.6.1. Modelo Longitud–Peso . . . . .	25
3.6.2. Modelo de crecimiento o de Longitud–Edad . . . . .	28
3.6.3. Modelo Stock–Reclutamiento de Ricker . . . . .	30

3.6.4. Modelo estadístico de evaluación del stock basado en captura por edad y por tallas . . . . .	32
<b>4. Materiales y Métodos</b>	<b>34</b>
4.1. Datos . . . . .	34
4.1.1. Datos simulados . . . . .	35
4.1.2. Datos observados . . . . .	35
4.2. Software . . . . .	36
4.3. Modelos matemáticos en dinámica poblacional . . . . .	37
4.3.1. Modelo Longitud–Peso . . . . .	37
4.3.2. Longitud–Edad (von Bertalanffy) . . . . .	37
4.3.3. Stock–Reclutamiento (Ricker) . . . . .	37
4.4. Metodología . . . . .	37
4.5. Comparación de métodos de estimación . . . . .	38
4.6. Comparación de algoritmos de optimización . . . . .	39
4.7. Aplicación de métodos de estimación y optimización a la cali- bración de un modelo de dinámica poblacional del jurel ( <i>Tra- churus murphyi</i> ) frente a Perú . . . . .	40
<b>5. Resultados y Discusión</b>	<b>42</b>
5.1. Resultados . . . . .	42
5.1.1. Simulación de Montecarlo . . . . .	42
5.1.2. Longitud–peso . . . . .	45
5.1.3. Longitud–edad . . . . .	45
5.1.4. Stock-reclutamiento . . . . .	50
5.1.5. Modelo estadístico de evaluación del stock basado en captura por edad y por tallas . . . . .	50
5.2. Discusión . . . . .	56
5.3. Perspectivas . . . . .	59
<b>Bibliografía</b>	<b>60</b>

# Índice de figuras

3.1. Problema Directo: Predicción de datos basándose en un modelo.	5
3.2. Problema Inverso: Estimación de parámetros. . . . .	6
5.1. Simulaciones Montecarlo de los parámetros para el modelo Longitud–Peso. . . . .	43
5.2. Simulaciones Montecarlo de los parámetros para el modelo von Bertalanffy. . . . .	44
5.3. Resultados de las estimaciones de los parámetros del modelo Longitud-Peso, para datos con diferente grado de dispersión. La línea discontinua roja indica el valor real del parámetro. . .	46
(a). Resultados de las simulaciones Montecarlo con 100 repeticiones (n=100) . . . . .	46
(b). Resultados de las simulaciones Montecarlo con 1000 repeticiones (n=1000) . . . . .	46
(a). Resultados de las simulaciones Montecarlo con 100 repeticiones (n=100) . . . . .	46
(b). Resultados de las simulaciones Montecarlo con 1000 repeticiones (n=1000) . . . . .	46
5.4. Resultados de la solución analítica del modelo Longitud–Peso.	47
5.5. Resultados de las estimaciones de los parámetros del modelo Von Bertalanffy, para datos con diferente grado de dispersión .	48
5.6. Resultados de las estimaciones de los parámetros del modelo Stock-Reclutamiento, para datos con diferente grado de dispersión . . . . .	51
5.7. Desembarque anual de jurel durante el periodo 1970-2012 . . .	52
5.8. Biomasa acústica estimada durante el periodo 1970-2012 . . .	53

5.9. Captura por unidad de esfuerzo (CPUE) estandarizada durante el periodo 2002-2012 . . . . .	53
5.10. Estimación del stock de biomasa desovante del jurel (1970-2012) calculado con los tres modelos estudiados . . . . .	54
5.11. Estimación del número de reclutas para el jurel (1970-2012) . . . . .	54
5.12. Estimación de la biomasa total de jurel (1970-2012) . . . . .	55
5.13. Estimación de la mortalidad por pesca del jurel (1970-2012) . . . . .	55
5.14. Estimación del número de individuos de jurel (1970-2012) . . . . .	56

# Índice de cuadros

4.1. Métodos de estimación utilizados para cada modelo. . . . .	39
4.2. Tipo de derivada para cada método de optimación y software donde se encuentra implementado. . . . .	39
5.1. Cuadro comparativo de los resultados obtenidos por cada mé- todo utilizado en el modelo Longitud–Peso. . . . .	49
5.2. Cuadro comparativo de los resultados obtenidos por cada mé- todo utilizado en el modelo Von Bertalanffy. . . . .	49
5.3. Cuadro comparativo de los resultados obtenidos por cada mé- todo utilizado en el modelo Stock–Reclutamiento. . . . .	50

# Capítulo 1

## Introducción

En modelamiento matemático, la estimación de parámetros a partir de datos observados, se denomina problemas inversos. Hadamard (1902) propuso que los modelos de fenómenos físicos deberían cumplir con tener solución, la cual además debería ser única y depender continuamente de los datos observados. Sin embargo, cuando se trabaja con modelos físicos (o biológicos), los problemas inversos son generalmente mal planteados, de manera que la solución puede no existir, no ser única o no ser estable. La razón podría ser la existencia de soluciones fuera del espacio de parámetros con sentido físico o a que las relaciones no lineales en los modelos introducen compensaciones entre los parámetros proporcionando muchas soluciones aceptables. Adicionalmente, la discretización que suele realizarse para encontrar soluciones numéricas puede conducir también a inestabilidades numéricas que hacen que este tipo de problemas sean además mal condicionados. Sin embargo, es posible introducir información adicional (e.g. restricción del espacio de parámetros, suavizamiento, penalizaciones por incremento de complejidad o número de parámetros) que permita convertir el problema inverso en bien planteado. Estos procedimientos, denominados *regularización*, permiten adicionalmente prevenir el riesgo de sobreparametrización en los modelos. Sin embargo, a pesar de lograr definir un problema bien planteado, es posible no encontrar la solución óptima si no se utiliza un algoritmo de optimización eficiente para el problema en particular. Esto es particularmente un problema cuando software tipo caja negra son usados para la optimización y el problema no cumple con las condiciones que necesita el algoritmo para asegurar la convergencia a la solución óptima. En este sentido, es importante analizar

las características de los problemas inversos impuestos por la estimación de parámetros en modelos de dinámica de poblaciones, de manera que se pueda tener mayor confianza en los parámetros estimados, los cuales suelen tener una interpretación biológica directa y son importantes para la evaluación y el manejo de recursos pesqueros.

# Capítulo 2

## Objetivos

### 2.1. Objetivo general

Comparar la aplicabilidad y eficiencia de diferentes algoritmos de optimización para la solución de problemas inversos en modelos de dinámica de poblaciones.

### 2.2. Objetivos específicos

1. Implementación computacional de un modelo de dinámica poblacional de un stock marino en AD Model Builder.
2. Comparación de eficiencia de diferentes algoritmos de optimización (algoritmos clásicos y algoritmos evolutivos).

# Capítulo 3

## Revisión de Literatura

### 3.1. Modelo matemático

Un modelo es la descripción matemática de un proceso real (Walter & Pronzato, 1997). Sirve para entender a profundidad un fenómeno (físico, químico o biológico). Sirve para estimar, a partir de medidas indirectas, cantidades para las cuales no se tiene disponible información, y para predecir comportamientos a corto o largo plazo.

Un modelo matemático es la relación entre datos y parámetros. Es una representación idealizada de un fenómeno tal que las desventajas de la abstracción excesiva es balanceada con una sobreabundancia de detalles (Gallucci & Quinn, 1979), y puede ser representado como:

$$G(m) = d, \tag{3.1}$$

donde  $G$  es el operador que relaciona el modelo  $m$  y el conjunto de datos observados  $d$  (Aster et al., 2005).

En modelamiento matemático de sistemas físicos y biológicos existen dos formas de estudiar un fenómeno, mediante el problema directo y el problema inverso.

#### 3.1.1. Problema Directo

Resolver un problema directo significa predecir los valores libres de error de los parámetros observados,  $d$  (3.2), que deberían corresponder a un modelo

dado,  $m$ .

$$m \mapsto d = G(m) \quad (3.2)$$

donde  $d = G(m)$  es una notación corta para el conjunto de ecuaciones  $d_i = G_i(m_1, m_2, \dots)$  para  $i = 1, 2, \dots$ .

El operador  $G$ , usualmente no lineal, es llamado el *operador directo*.

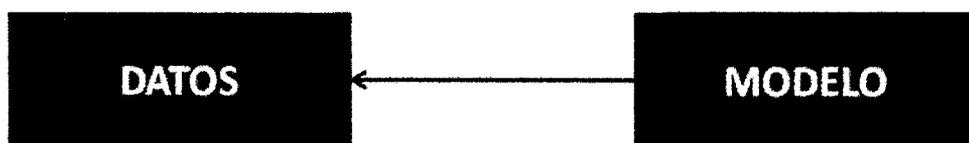


Figura 3.1: Problema Directo: Predicción de datos basándose en un modelo.

### 3.1.2. Problema Inverso

La teoría del problema inverso es la teoría matemática que describe cómo la información de un sistema parametrizado (físico o biológico) puede ser derivado de los datos de observación, relaciones teóricas entre los parámetros del modelo y los datos, y la información previa (Tarantola, 2005; Tarantola, 2006).

En ese sentido, los métodos inversos están recibiendo mayor atención en ecología marina (Dowd & Meyer, 2003).

El problema inverso consiste en utilizar el resultado real de algunas medidas para deducir el valor de los parámetros que caracterizan el sistema (Monett & Burkhard, 2004).

El problema inverso contrasta con el problema directo, donde se predicen los datos a partir de los parámetros y de un modelo. Normalmente el problema inverso es más difícil de resolver que su correspondiente problema directo (Gasulla Forner, 1999).

Menke (1989) dice que el problema inverso es simplemente el conjunto de métodos usados para extraer información útil de nuestro entorno a partir de medidas físicas o datos. La información útil vendrá especificada como valores numéricos de alguna propiedad de este entorno. Estas propiedades también se referirán como parámetros del modelo. Se presupone que hay algún método específico (normalmente una teoría matemática o modelo) que

relaciona los parámetros con los datos, aunque por lo general, este no es tan evidente y/o está encapsulado en algún software, que por su uso cotidiano, se terminan usando como *caja negra*, olvidando aspectos esenciales del proceso de optimización (como por ejemplo, verificar si los datos ingresados al modelo cumplen con las condiciones mínimas de uso).



Figura 3.2: Problema Inverso: Estimación de parámetros.

Para el estudio de dicha *caja negra*, estudiamos los fundamentos del proceso de optimización, así como la elección del método de optimización y el método de estimación.

## 3.2. Fundamentos de optimización sin restricciones

En optimización sin restricciones se minimiza una función objetivo que depende de variables reales irrestrictas sobre todo el dominio de la función. La formulación matemática es:

$$\min_{x \in \mathbb{R}^n} f(x) \quad (3.3)$$

donde  $x \in \mathbb{R}^n$  es un vector real con  $n \geq 1$  y  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es una función suave (suficientemente regular).

**Definición 3.2.1.** *Se dice que una función  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es continuamente diferenciable cuando es diferenciable, con derivada continua. A las funciones continuamente diferenciables también se les llama de clase  $C^1$  o suaves.*

*Se dice que una función  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es dos veces continuamente diferenciable o de clase  $C^2$  cuando es dos veces diferenciable, con derivadas continuas.*

**Definición 3.2.2.** Un punto  $x^*$  es un mínimo global si  $f(x^*) \leq f(x)$  para todo  $x \in \mathbb{R}^n$ . Este punto es un mínimo local si existe una vecindad  $\mathcal{V}$  de  $x^*$  tal que  $f(x^*) \leq f(x)$  para todo  $x \in \mathcal{V}$ . Análogamente, es un mínimo local estricto si existe una vecindad  $\mathcal{V}$  de  $x^*$  tal que  $f(x^*) < f(x)$  para todo  $x \in \mathcal{V} \setminus \{x^*\}$ .

De la definición dada anteriormente, pareciera que la única manera de saber si un vector  $x^*$  es mínimo local es examinando todos los puntos en las inmediaciones de este. Si la función  $f$  es suficientemente regular, existen modos eficientes y prácticos de identificar mínimos locales. En particular, si  $f$  es dos veces diferenciable, puede decirse si  $x$  es un mínimo local examinando su gradiente  $\nabla f(x)$  y su hessiano  $\nabla^2 f(x)$ .

**Teorema 3.2.1.** (Teorema de Taylor) Sea  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  continuamente diferenciable y  $p \in \mathbb{R}^n$ . Entonces tenemos que

$$f(x + p) = f(x) + \nabla f(x + tp)^\top p \quad (3.4)$$

para algún  $t \in \langle 0, 1 \rangle$ . Más aún, si  $f$  es dos veces continuamente diferenciable, se tiene que

$$\nabla f(x + p) = \nabla f(x) + \int_0^1 \nabla^2 f(x + tp) p dt, \quad (3.5)$$

y que

$$f(x + p) = f(x) + \nabla f(x)^\top p + \frac{1}{2} p^\top \nabla^2 f(x + tp) p, \quad (3.6)$$

para algún  $t \in \langle 0, 1 \rangle$ .

Las condiciones necesarias de optimalidad se derivan de asumir que  $x^*$  es un mínimo local y luego probar ciertas condiciones sobre  $\nabla f(x^*)$  y  $\nabla^2 f(x^*)$ .

**Teorema 3.2.2.** (Condición necesaria de primer orden) Si  $x^*$  es un mínimo local y  $f$  es continuamente diferenciable en una vecindad de  $x^*$ , entonces  $\nabla f(x^*) = 0$ .

**Definición 3.2.3.** Decimos que  $x^*$  es un punto estacionario si  $\nabla f(x^*) = 0$ .

**Definición 3.2.4.** Decimos que la matriz  $B$  es definida positiva si  $p^\top B p > 0$  para todo  $p \neq 0$ , y semidefinida positiva si  $p^\top B p \geq 0$  para todo  $p$ .

**Teorema 3.2.3.** (Condición necesaria de segundo orden) Si  $x^*$  es un mínimo local de  $f$  y  $\nabla^2 f$  existe y es continua en una vecindad de  $x^*$ , entonces  $\nabla f(x^*) = 0$  y  $\nabla^2 f(x^*)$  es semidefinida positiva.

**Teorema 3.2.4.** (Condiciones suficientes de segundo orden) *Supongamos que  $\nabla^2 f$  es continua en una vecindad de  $x^*$  y que  $\nabla f(x^*) = 0$  y  $\nabla^2 f(x^*)$  es definida positiva. Entonces  $x^*$  es un mínimo local estricto de  $f$ .*

### 3.3. Métodos de optimización

El método de optimización es la parte fundamental de todo el proceso de optimización. Es la estrategia que se considera al abordar un problema. Tiene como finalidad la minimización de una función objetivo.

Existen diversos métodos de optimización, dependiendo de las condiciones de la función objetivo y los datos ingresados al modelo. Dichos métodos se pueden clasificar en dos grandes grupos, los basados en derivadas y los libres de éstas. En cada uno de estos grupos, existen además condiciones que se pueden considerar al momento de la elección de algún método para la realización de un modelo.

Para este estudio nos centramos en cinco métodos, los métodos basados en derivadas como el método *BFGS*, el método *L-BFGS-B* y el método de *gradiente conjugado*, métodos deterministas como el método *Nelder-Mead*, y el método libre de derivadas *diferenciación automática*.

#### 3.3.1. Métodos basados en derivadas

Sea  $f : C \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  una función dada. Supongamos que  $f$  es de clase  $C^1$  sobre  $C \subseteq \mathbb{R}^n$ . Consideremos el problema sin restricciones

$$\text{mín } f(x), \quad x \in C \tag{3.7}$$

Dado un punto inicial  $x_0 \in C$ , aproximación de la solución del problema (3.7). Construimos una sucesión  $\{x_k\} \subset C$  tal que  $f(x_{k+1}) < f(x_k)$  para todo  $k$  hasta que una condición de optimalidad sea verificada (Izmailov & Solodov, 2007; Crouzeix et al., 2011).

Para un punto  $x_k \in C$  conocido y disponible,

- Se determina una dirección de descenso  $d_k \in C$  tal que  $\langle \nabla f(x_k), d_k \rangle < 0$  y se conoce la existencia de  $\hat{t}_k > 0$ , con  $x_k + td_k \in C$  para todo  $t \in [0, \hat{t}_k]$ . Si  $C = \mathbb{R}^n$ , entonces  $\hat{t}_k = +\infty$ .

- Se busca  $t_k \in ]0, \hat{t}_k]$  tal que  $f(x_k + t_k d_k) < f(x_k)$ .
- Se hace  $x_{k+1} = x_k + t_k d_k$  y se repite el proceso.

**Definición 3.3.1.** *A este último procedimiento se le conoce como búsqueda lineal.*

**Definición 3.3.2.** *Decimos que  $d \in \mathbb{R}^n$  es una dirección de descenso de la función  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  en el punto  $x \in \mathbb{R}^n$ , si existe  $\varepsilon > 0$  tal que*

$$f(x + td) < f(x) \quad \forall t \in ]0, \varepsilon]. \quad (3.8)$$

**Lema 3.3.1.** (Direcciones de descenso) *Sea  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  una función diferenciable en el punto  $x \in \mathbb{R}^n$ . Entonces:*

- (a) *Para toda dirección de descenso  $d$  se tiene que  $\langle \nabla f(x), d \rangle \leq 0$ .*
- (b) *Si  $d \in \mathbb{R}^n$  satisface  $\langle \nabla f(x), d \rangle < 0$ , entonces  $d$  es una dirección de descenso.*

*Demostración.* Sea  $d$  dirección de descenso. Para todo  $t > 0$  suficientemente pequeño, por la diferenciabilidad de  $f$  en  $x$

$$\begin{aligned} 0 > f(x + td) - f(x) &= t \langle \nabla f(x), d \rangle + o(t) \\ &= t \left( \langle \nabla f(x), d \rangle + \frac{o(t)}{t} \right) \end{aligned} \quad (3.9)$$

Dividiendo los dos lados de la desigualdad por  $t > 0$  y tomando el límite cuando  $t \rightarrow 0^+$ , obtenemos  $0 \geq \langle \nabla f(x), d \rangle$ , por lo que el ítem (a) queda probado.

Supongamos ahora que  $\langle \nabla f(x), d \rangle < 0$ . Tenemos que

$$f(x + td) - f(x) = t \left( \langle \nabla f(x), d \rangle + \frac{o(t)}{t} \right). \quad (3.10)$$

En particular, para todo  $t > 0$  suficientemente pequeño tenemos

$$\langle \nabla f(x), d \rangle + \frac{o(t)}{t} \leq \frac{1}{2} \langle \nabla f(x), d \rangle < 0, \quad (3.11)$$

lo que implica que  $f(x + td) - f(x) < 0$ , por lo que, por definición,  $d$  es dirección de descenso.  $\square$

**Lema 3.3.2.** Si la función  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es dos veces diferenciable en el punto  $x \in \mathbb{R}^n$ , entonces

(a) Para toda dirección de descenso  $d$  tal que  $\langle \nabla f(x), d \rangle = 0$ , se tiene que  $\langle \nabla^2 f(x)d, d \rangle \leq 0$ .

(b) Si  $d \in \mathbb{R}^n$  satisface  $\langle \nabla f(x), d \rangle = 0$  y  $\langle \nabla^2 f(x)d, d \rangle < 0$ , entonces  $d$  es una dirección de descenso.

*Demostración.* Sea  $d$  dirección de descenso. Para  $t > 0$  suficientemente pequeño, por la doble diferenciable de  $f$  en  $x$ , tenemos

$$0 > f(x + td) - f(x) = t\langle \nabla f(x), d \rangle + \frac{1}{2}t^2\langle \nabla^2 f(x)d, d \rangle + o(t^2) \quad (3.12)$$

Como  $\langle \nabla f(x), d \rangle = 0$ , la expresión (3.12) se reduce a

$$0 > \frac{1}{2}t^2\langle \nabla^2 f(x)d, d \rangle + o(t^2) \quad (3.13)$$

Dividiendo los dos lados de la desigualdad por  $\frac{t^2}{2} > 0$  y tomando el límite cuando  $t \rightarrow 0^+$ , obtenemos  $\langle \nabla^2 f(x)d, d \rangle \leq 0$ , por lo que queda demostrado el ítem (a).

Sea  $d \in \mathbb{R}^n$  tal que  $\langle \nabla f(x), d \rangle = 0$  y  $\langle \nabla^2 f(x)d, d \rangle < 0$ . Luego, tenemos

$$\begin{aligned} f(x + td) - f(x) &= \frac{1}{2}t^2\langle \nabla^2 f(x)d, d \rangle + o(t^2) \\ &= t^2 \left( \frac{1}{2}\langle \nabla^2 f(x)d, d \rangle + \frac{o(t^2)}{t^2} \right) \end{aligned} \quad (3.14)$$

En particular, para todo  $t > 0$  suficientemente pequeño tenemos

$$\frac{1}{2}\langle \nabla^2 f(x)d, d \rangle + \frac{o(t^2)}{t^2} \leq \frac{1}{4}\langle \nabla^2 f(x)d, d \rangle < 0, \quad (3.15)$$

lo que implica que  $f(x + td) - f(x) < 0$ , por lo que, por definición,  $d$  es dirección de descenso.  $\square$

Por lo que el esquema iterativo general de los métodos de descenso consiste en:

$$x_{k+1} = x_k + \alpha_k d_k, \quad k = 0, 1, \dots, \quad (3.16)$$

donde  $d_k$  es la dirección de descenso y  $\alpha_k > 0$ .

La elección más obvia de una dirección de descenso  $d_k$  de la función  $f$  en el punto  $x_k \in \mathbb{R}^n$  es  $d_k = -\nabla f(x_k)$  (asumiendo que  $\nabla f(x_k) \neq 0$ ), aunque frecuentemente, los métodos correspondientes a esta elección poseen convergencia muy lenta.

Desde el punto de vista práctico, más importantes son los métodos con  $d_k = -Q_k \nabla f(x_k)$ , donde  $Q_k \in \mathbb{R}^{n \times n}$  es una matriz simétrica definida positiva.

Si se escoge  $Q_k = I$  (matriz identidad), se obtiene  $d_k = -\nabla f(x_k)$ .

Por el lema 3.3.1, si  $\nabla f(x_k) \neq 0$ ,  $d_k = -Q_k \nabla f(x_k)$  también es una dirección de descenso de  $f$  en  $x_k$  y el análisis teórico de convergencia global no es muy diferente del análisis para el caso de  $d_k = -\nabla f(x_k)$ .

### 3.3.1.1. Método de Newton

Una iteración del método de Newton es numéricamente más costoso en comparación con el método del gradiente. Los métodos quasi-Newton son ligeramente más lentos que el método de Newton, pero tiene iteraciones mucho más costosas, en particular, no tanto como los métodos del gradiente (Izmailov & Solodov, 2007).

Sea  $f$  diferenciable,  $\nabla^2 f(x)$  no es conocido, por lo que se utilizan aproximaciones.

Consideremos el problema de optimización

$$\min f(x), \quad x \in \mathbb{R}^n \quad (3.17)$$

donde  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es una función dos veces diferenciable en  $\mathbb{R}^n$ .

El método calcula una aproximación de  $\nabla^2 f(x)$ . El método consiste en tomar  $d_k$  como la solución del sistema lineal

$$\nabla^2 f(x_k) d_k = -\nabla f(x_k). \quad (3.18)$$

Así, el método de Newton consiste de la iteración

$$x_{k+1} = x_k - \alpha_k (\nabla^2 f(x_k))^{-1} \nabla f(x_k), \quad (3.19)$$

asumiendo que la dirección de Newton

$$d_k = -(\nabla^2 f(x_k))^{-1} \nabla f(x_k) \quad (3.20)$$

es definida y es una dirección de descenso (Bertsekas, 1999; Lange, 2004; Fletcher, 2000).

Asumamos que  $f \in \mathcal{C}^2$  y que la matriz Hessiana satisface la condición de Lipschitz  $\|G(x) - G(y)\| \leq \lambda\|x - y\|$  en una vecindad de un minimizador local  $x^*$ . Si  $x^{(k)}$  es suficientemente cercano a  $x^*$  para algún  $k$ , y si  $G^*$  es definida positiva, entonces el método de Newton está bien definido para todo  $k$ , y converge al segundo orden.

### 3.3.1.2. Método Quasi-Newton

Los métodos Quasi-Newton sólo requieren el gradiente de la función objetivo en cada iteración. Midiendo los cambios de los gradientes, construyen un modelo de la función objetivo que es lo suficientemente bueno para producir convergencia superlineal.

Evita el cálculo de las hessianas y por lo tanto, también las resolución de los sistemas lineales.

Consiste en tomar como dirección de descenso en la iteración  $k$

$$d_k = -Q_k \nabla f(x_k) \quad (3.21)$$

tal que

$$Q_{k+1} = Q_k + \left(1 + \frac{\langle y_k, Q_k y_k \rangle}{\langle y_k, \sigma_k \rangle}\right) \frac{\sigma_k \sigma_k^\top}{\langle y_k, \sigma_k \rangle} - \frac{\sigma_k y_k^\top Q_k + Q_k y_k \sigma_k^\top}{\langle y_k, \sigma_k \rangle} \quad (3.22)$$

con

$$\sigma_k = x_{k+1} - x_k \quad \text{y} \quad y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$$

donde  $Q_0$  es una matriz simétrica definida positiva, en particular  $Q_0 = I$ .

Se deduce fácilmente que, para todo  $k$ ,  $Q_k$  es simétrica definida positiva, dado que  $Q_0$  es simétrica definida positiva.

**Teorema 3.3.1.** *Si  $Q_k$  es simétrica definida positiva, definida como en (3.22), entonces  $Q_{k+1}$  es simétrica definida positiva.*

Es necesario que  $f$  sea de clase  $\mathcal{C}^2$  y que la matriz  $\nabla^2 f(x)$  sea definida positiva a fin que las matrices  $Q_k$  resulten simétricas definidas positivas y obtener la convergencia (Crouzeix et al., 2011).

Sea:

$$x^{k+1} = x^k + \alpha_k d_k, \quad k = 0, 1, \dots, \quad (3.23)$$

donde, para todo  $k$ ,  $Q_k \in \mathbb{R}^{n \times n}$  matriz simétrica definida positiva, y  $\alpha_k > 0$  es la longitud de paso calculado utilizando alguna de las reglas de búsqueda.

Para todo  $k$ , definimos

$$r^k = x^{k+1} - x^k \quad s^k = \nabla f(x^{k+1}) - \nabla f(x^k) \quad (3.24)$$

$$Q_{k+1}s^k = r^k \quad (3.25)$$

Davidon-Fletcher-Powell (DFP):

$$Q_{k+1} = Q_k + \frac{r^k(r^k)^\top}{\langle r^k, s^k \rangle} - \frac{(Q_k s^k)(Q_k s^k)^\top}{\langle Q_k s^k, s^k \rangle} \quad (3.26)$$

Broyden-Fletcher-Goldfarb-Shanno (BFGS):

$$Q_{k+1} = Q_k + \frac{(r^k - Q_k s^k)(r^k)^\top + r^k(r^k - Q_k s^k)^\top}{\langle r^k, s^k \rangle} - \frac{\langle r^k - Q_k s^k, s^k \rangle r^k (r^k)^\top}{\langle r^k, s^k \rangle^2} \quad (3.27)$$

considerando  $Q_0$  como la matriz identidad,  $Q_0 = I$ .

Los métodos quasi-newton son métodos de la forma (Izmailov & Solodov, 2007) (por el esquema iterativo)

$$x^{k+1} = x^k + \alpha_k d^k, \quad d^k = -Q_k f'(x^k), \quad k = 0, 1, \dots, \quad (3.28)$$

donde, para todo  $k$ ,  $Q_k \in \mathbb{R}^{n \times n}$  es una matriz simétrica definida positiva, y  $\alpha_k > 0$  es la longitud del paso calculado usando una regla de búsqueda lineal.

## Método BFGS

El algoritmo Quasi-Newton más popular es el método BFGS, el cual debe su nombre a sus descubridores Broyden, Fletcher, Goldfarb y Shanno, (Broyden, 1970a,9; Fletcher, 1970; Goldfarb, 1970; Shanno, 1970).

Empezamos la derivación formando el siguiente modelo cuadrático de la función objetivo en la iteración  $x_k$ :

$$m_k(p) = f_k + \nabla f_k^\top p + \frac{1}{2} p^\top B_k p \quad (3.29)$$

donde  $B_k \in \mathbb{R}^{n \times n}$  simétrica definida positiva que será actualizada en cada iteración.

El valor de la función y su gradiente en el punto  $p = 0$  son iguales a  $f_k$  y  $\nabla f_k$  respectivamente.

El minimizador  $p_k$  de este modelo convexo cuadrático, denotando por  $p_k = -B_k^{-1}\nabla f_k$ , es usado como la dirección de búsqueda, y la nueva iteración es

$$x_{k+1} = x_k + \alpha_k p_k \quad (3.30)$$

donde la longitud de paso  $\alpha_k$  es elegida para satisfacer las condiciones de Wolfe. Esta iteración es muy similar a la búsqueda lineal método de Newton; la principal diferencia es que es usada la hessiana aproximada  $B_k$  en lugar de la verdadera hessiana.

(Nash, 1990)

### Método L-BFGS-B

Es un algoritmo quasi-newton de memoria limitada para resolver problemas de optimización no lineal grandes con cotas (límites) simples en las variables. El problema tiene la forma general:

$$\begin{array}{l} \text{mín} \\ \text{sujeto a } l \leq x \leq u \end{array} f(x) \quad (3.31)$$

donde  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es una función no lineal cuya gradiente  $g$  es conocida, los vectores  $l$  y  $u$  representan el límite superior e inferior de las variables, y el número de variables  $n$  se asume muy grande.

El algoritmo no requiere segundas derivadas ni el conocimiento de la estructura de la función objetivo, y puede por lo tanto ser aplicada cuando la matriz hessiana no es fácil de calcular (Byrd et al., 1995). Una actualización quasi-Newton de memoria limitada es usada para aproximar la matriz hessiana, de tal manera que el almacenamiento requerido es lineal en  $n$ .

Se forma el siguiente modelo cuadrático de la función objetivo en la iteración  $x_k$ :

$$m_k(x) = f(x_k) + g_k^\top(x - x_k) + \frac{1}{2}(x - x_k)^\top B_k(x - x_k). \quad (3.32)$$

El algoritmo aproximadamente minimiza  $m_k(x)$  sujeto a los límites dados en (3.31).

La minimización se obtiene considerando el camino lineal por tramos

$$x(t) = P(x_k - tg_k, l, u), \quad (3.33)$$

obtenido por proyectar la dirección de máximo descenso en la región factible, donde:

$$P(x, l, u)_i := \begin{cases} l_i & \text{si } x_i < l_i \\ x_i & \text{si } x_i \in [l_i, u_i] \\ u_i & \text{si } x_i > u_i \end{cases} \quad (3.34)$$

### 3.3.1.3. Método de Gradiente Conjugado

Este método se usa tanto para resolver problemas de optimización lineal como no lineal (Nocedal & Wright, 2006).

#### Método de gradiente conjugado lineal

El método de gradiente conjugado lineal es un método iterativo para resolver un sistema lineal de ecuaciones

$$Ax = b, \quad (3.35)$$

donde  $A \in \mathbb{R}^{n \times n}$  matriz simétrica definida positiva y  $b \in \mathbb{R}^n$ . El problema puede expresarse equivalentemente como el problema de minimización

$$\text{mín } \phi(x) = \frac{1}{2}x^\top Ax - b^\top x, \quad (3.36)$$

pues tienen la misma solución única. De aquí, denotamos por

$$r(x) = \nabla \phi(x) = Ax - b \quad (3.37)$$

en particular, si  $x = x_k$  tenemos

$$r_k = Ax_k - b \quad (3.38)$$

**Método de las direcciones conjugadas** Sea el conjunto de vectores no nulos  $\{p_i\}_{i=1}^n$ , linealmente independientes y conjugados respecto a la matriz simétrica definida positiva  $A$ , esto es

$$p_k^\top A p_j = 0, \text{ si } k \neq j, \quad (3.39)$$

y dado un punto de inicio  $x_0 \in \mathbb{R}$ , generamos la sucesión  $\{x_k\}$  de la forma

$$x_{k+1} = x_k + \alpha_k p_k, \quad (3.40)$$

donde  $\alpha_k$  es el minimizador unidimensional de la función cuadrática  $\phi(\cdot)$  a lo largo de  $x_k + \alpha p_k$ , dado explícitamente por

$$\alpha_k = -\frac{r_k^\top p_k}{p_k^\top A p_k} \quad (3.41)$$

En el método de gradiente conjugado, cada dirección  $p_k$  se elige como una combinación lineal del residual negativo  $-r_k$  y la dirección previa  $p_{k-1}$ ,

$$p_k = -r_k + \beta_k p_{k-1}, \quad (3.42)$$

donde el escalar  $\beta_k$  es determinado por

$$\beta_k = \frac{r_k^\top A p_{k-1}}{p_{k-1}^\top A p_{k-1}}. \quad (3.43)$$

Debido a que el problema consiste en minimizar el residual  $r_k$  a través de la dirección  $p_k$ , se cumple la propiedad

$$r_k^\top p_i = 0, \quad \text{para } i = 0, 1, \dots, k-1. \quad (3.44)$$

De ahí, podemos obtener una forma más económica con

$$\alpha_k = \frac{r_k^\top r_k}{p_k^\top A p_k} \quad (3.45)$$

y

$$\beta_{k+1} = \frac{r_{k+1}^\top r_{k+1}}{r_k^\top r_k} \quad (3.46)$$

### Métodos de gradiente conjugado no lineal

Notamos en (3.36) que el método de gradiente conjugado lineal puede ser visto como la minimización de una función convexa cuadrática  $\phi$ . Se puede extender el método a funciones convexas generales, incluso a funciones generales no lineales  $f$ .

**El método Fletcher–Reeves** Fletcher y Reeves mostraron cómo extender el método de gradiente conjugado a funciones no lineales cambiando la longitud de paso  $\alpha_k$ , el cual minimiza  $\phi$  a través de la dirección de búsqueda  $p_k$ , tenemos que realizar una búsqueda de línea que identifique a un mínimo aproximado de la función no lineal  $f$  a lo largo de  $p_k$ . Además, el residual  $r$ , que es simplemente el gradiente de  $\phi$  en el caso lineal, debe ser reemplazado por el gradiente de la función objetivo no lineal  $f$ .

Se define la dirección de búsqueda  $p_k$  como

$$p_{k+1} = -\nabla f_{k+1} + \beta_{k+1}^{FR} p_k \quad (3.47)$$

donde el parámetro  $\beta_k$  está dado por

$$\beta_{k+1}^{FR} = \frac{\nabla f_{k+1}^\top \nabla f_{k+1}}{\nabla f_k^\top \nabla f_k}. \quad (3.48)$$

**El método Polak–Ribière** Esta variante del método de Fletcher–Reeves, propuesta por Polak y Ribière, define el parámetro  $\beta_k$  como

$$\beta_{k+1}^{PR} = \frac{\nabla f_{k+1}^\top (\nabla f_{k+1} - \nabla f_k)}{\|\nabla f_k\|^2}. \quad (3.49)$$

### 3.3.2. Métodos libres de derivadas

La caracterización matemática “estándar” de un mínimo local está dada por la condición necesaria de primer orden, la cual requiere para funciones continuamente diferenciales que la derivada de primer orden sea cero, (Conn et al., 2009).

Muchas aplicaciones prácticas requieren la optimización de funciones cuyas derivadas no son fáciles de obtener, para resolver este tipo de problemas, en principio, se puede aproximar el gradiente (y posiblemente el hessiano) usando diferencias finitas para luego proceder con dichas aproximaciones usando los algoritmos anteriores (Nocedal & Wright, 2006).

Hay situaciones en las que no pueden ser aplicadas las diferencias finitas ni la diferenciación automática.

### 3.3.2.1. Métodos deterministas

**Definición 3.3.3.** Un conjunto  $P \subset \mathbb{R}^n$  es un poliedro si hay un sistema de un número finito de desigualdades  $Ax < b$  tal que

$$P = \{x \in \mathbb{R}^n \mid Ax < b\}. \quad (3.50)$$

**Definición 3.3.4.** Un conjunto  $P \subset \mathbb{R}^n$  es un polígono si es un poliedro acotado.

### Nelder-Mead

Este método numérico minimiza una función objetivo  $f$  en un espacio multidimensional (Nelder & Mead, 1965).

Sea un polígono de  $n+1$  vértices en un espacio  $n$ -dimensional, sean  $\{P_i\}_{i=0}^n$  dichos vértices con valores  $y_i = f(P_i)$ ,  $i = 0, 1, \dots, n$ . Denotamos los puntos  $P_l$  y  $P_h$  asociados a  $y_l = f(P_l)$  y  $y_h = f(P_h)$  respectivamente, tales que  $y_l = \max_i(y_i)$  y  $y_h = \min_i(y_i)$ ; el punto  $\bar{P} = \text{cent}\{P_i\}_{i=0}^n = \sum_{i=0}^n \frac{P_i}{n}$ , con  $i \neq h$ , centroide de los puntos con  $i \neq h$ ; y denotemos la distancia de  $P_i$  a  $P_j$  como  $[P_i P_j]$ .

En el algoritmo,  $P_h$  es reemplazado por un nuevo punto en cada iteración a través de tres operaciones definidas, tales como reflexión, contracción y expansión.

Primero se define el punto  $P^*$  por

$$P^* = (1 + \alpha)\bar{P} - \alpha P_h \quad (3.51)$$

donde  $\alpha > 0$  llamada coeficiente de reflexión. Así,  $P^*$  está en la línea que une  $\bar{P}$  y  $P_h$ , en el lado opuesto a  $P_h$  respecto a  $\bar{P}$ , y se cumple  $[P^* \bar{P}] = \alpha[P_h \bar{P}]$ .

Si  $y_l < y^* < y_h$ ,  $P_h$  es reemplazado por  $P^*$  y se realiza nuevamente el proceso. Si  $y^* < y_l$  (si la reflexión produce un nuevo mínimo), entonces expandimos  $P^*$  al punto  $P^{**}$  definido por

$$P^{**} = \gamma P^* + (1 - \gamma)\bar{P} \quad (3.52)$$

con  $\gamma > 1$  la relación de la distancia  $[P^{**} \bar{P}]$  a  $[P^* \bar{P}]$ , llamado coeficiente de expansión.

Si  $y^{**} < y_i$  reemplazamos  $P_h$  por  $P^{**}$  y volvemos a realizar el proceso. Si  $y^{**} > y_i$  entonces tenemos lo que se llama “expansión fallida”, y reemplazamos  $P_h$  por  $P^*$  antes de volver a realizar el proceso.

Si al reflejar  $P$  a  $P^*$  tenemos que  $y^* > y_i$ , para todo  $i \neq h$ , es decir, al reemplazar  $P$  por  $P^{**}$  tenemos que  $y^{**}$  es el máximo, entonces definimos un nuevo  $P_h$ , el cual será el punto  $P_h$  o  $P^*$  anterior (el de menor valor  $y$ ), y determinamos

$$P^{**} = \beta P_h + (1 - \beta) \bar{P} \quad (3.53)$$

donde  $\beta \in \langle 0, 1 \rangle$  es llamado coeficiente de contracción y es la relación de la distancia  $[P^{**} \bar{P}]$  a  $[P \bar{P}]$ . Reemplazamos  $P_h$  por  $P^{**}$  y volvemos a realizar el proceso, a menos que  $y^{**} > \min(y_h, y^{**})$ , es decir, el punto contraído es peor que el mejor de  $P_h$  y  $P^*$ . Para este caso de contracción fallida, reemplazamos todos los puntos  $P_i$  por  $\frac{P_i + P_i}{2}$  y volvemos a realizar el proceso.

Una expansión fallida puede deberse como resultado de entrar en un valle ( $P^*$ ) pero en un ángulo tal que  $P^{**}$  está más arriba, al lado opuesto.

Una contracción fallida es mucho más rara, puede ocurrir cuando un valle es curvo y un punto del simplex está más alejado del fondo del valle que los otros, la contracción puede entonces causar que el punto reflejado se mueva lejos del fondo del valle en lugar de acercarse hacia él.

Las acciones propuestas contraen el simplex hacia el punto más bajo y, eventualmente, llevarán todos los puntos hacia el valle.

Los coeficientes  $\alpha, \beta, \gamma$  son los factores por el cual el volumen del simplex es cambiado al realizar las operaciones de reflexión, contracción y expansión respectivamente.

El algoritmo de Nelder-Mead es un método de búsqueda directa en el sentido de que la función objetivo es evaluada en un número finito de puntos por iteración, y decide que acción tomar después basada únicamente en los valores de la función y sin ninguna aproximación explícita o implícita de la derivada o construcción de modelo, (Conn et al., 2009).

### 3.3.2.2. Métodos estocásticos

La diferenciación automática es el nombre genérico de las técnicas que utilizan la representación computacional de una función para producir valores analíticos de las derivadas. Algunas técnicas producen el código para las derivadas en un punto genérico  $x$  manipulando el código de la función

directamente. Otras técnicas mantienen un registro de los cálculos realizados durante la evaluación de la función en un punto específico  $x$  y luego revisan esta información para producir una serie de derivadas en  $x$ .

Las técnicas de diferenciación automática se basan en la observación de que cualquier función, no importa lo complicada que sea, se evalúa mediante la realización de una serie de operaciones elementales simples de sólo uno o dos argumentos a la vez. Las operaciones de dos argumentos incluyen la suma, multiplicación, división, y potencia  $a^b$ . Ejemplos de operaciones de un sólo argumento incluyen las funciones trigonométricas, exponenciales y logarítmicas. Otro elemento común de las diferentes herramientas de diferenciación automática es el uso de la regla de la cadena. Esta es la famosa regla de cálculo elemental que dice que si  $h$  es una función del vector  $y \in \mathbb{R}^m$ , que a su vez es una función del vector  $x \in \mathbb{R}^n$ , podemos escribir la derivada de  $h$  con respecto a  $x$  de la siguiente manera:

$$\nabla_x h(y(x)) = \sum_{i=1}^m \frac{\partial h}{\partial y_i} \nabla y_i(x). \quad (3.54)$$

Las derivadas juegan un rol importante en análisis de sensibilidad (validación del modelo), problemas inversos (asimilación de datos), y diseño de optimización (simulación para la elección de parámetros) (Griewank & Walther, 2008).

Los métodos eficientes para optimizar la función objetivo están basadas en derivadas, las cuales típicamente son aproximaciones numéricas usando diferencias finitas, o alternativamente, evaluando una expresión analítica de las derivadas.

El método anterior puede ser inestable, y es ineficiente cuando hay muchos parámetros. Por ello, se prefieren las derivadas analíticas, pero éstas pueden ser difíciles de calcular para modelos complejos implementados en cientos de líneas de código (computacional). Una tercera alternativa es proporcionada por la diferenciación automática (AD por sus siglas en inglés), la cual es una técnica para la evaluación numérica de derivadas de una función matemática disponible en el formulario de un programa computacional.

AD Model Builder (ADMB), un sistema basado en AD (diferenciación automática) para hacer estimaciones (estadísticas) de parámetros, que incluye un conjunto completo de herramientas para desarrollar modelos estadísticos

no lineales. ADMB integra a la perfección AD con una función minimizante, la cual es usada para minimizar una función objetivo.

**Implementación de AD** Existen dos modos básicos de diferenciación automática: el modo directo (*forward*) y el modo inverso (*reverse*).

En el modo directo, las derivadas de las variables temporales con respecto a las variables independientes se propagan en paralelo con el flujo del problema general. Este tiene una deficiencia, pues el costo de propagar las derivadas crece en proporción al número de las variables independientes.

El modo inverso consiste en evaluar la función objetivo, guardando en memoria el valor de cada cantidad intermedia, esto es conocido como “barrido directo”, en el cual ninguna derivada es calculada. El principal objetivo es encontrar la derivada de la variable de salida con respecto a cada variable intermedia, resultando en una secuencia de ‘sensibilidades’. Cada paso implica la regla de la cadena, y el orden inverso de los cálculos es esencial.

Para el modo inverso, el gradiente de la función objetivo puede ser obtenida a un costo no mayor a cuatro veces el costo de evaluar la función objetivo (Griewank & Walther, 2008). La desventaja del modo inverso es la necesidad de almacenar en memoria los valores de las variables temporales calculadas durante el barrido directo. ADMB usa el modo inverso como estrategia general para calcular las derivadas de primer orden.

Para programas que contienen bucles que despliegan en largas listas, el grafo computacional puede convertirse tan grande para ajustar(?) en la memoria física, que por ello tiene que ser escrita en un archivo.

Escribirlo a un archivo hace la ejecución del programa muy lento. Para contrarrestar ese fenómeno, ADMB implementa el modo inverso de AD por sobrecarga de operadores en C++. Con esta estrategia la regla de la cadena es aplicada al tiempo de ejecución (del programa ejecutable), en oposición(en vez de) a generar el código de derivadas al momento de compilar. Esta técnica es conocida como la “transformación de código” (modificación del código fuente).

La parte interna de ADMB que sobrecarga todos los operadores matemáticos y funciones en C++ y establece las estructuras de los datos necesarios para el barrido directo, es la librería conocida como AUTODIF.

La librería AUTODIF y ADMB fueron inicialmente desarrollados para

implementar modelos estadísticos de evaluación pesquera integrados altamente parametrizados que eran esencialmente imposibles de estimar con el software disponible en ese tiempo (Fournier et al., 2012).

### 3.4. Métodos de estimación

La estimación de parámetros conlleva encontrar valores para los parámetros que provean el mejor ajuste entre el modelo y los datos disponibles, de acuerdo al criterio que determine la bondad del ajuste a los datos de una particular combinación de modelo y parámetros estimados.

Existen dos criterios ampliamente utilizados para la bondad de ajuste. El más comúnmente utilizado es el método de mínimos cuadrados, en el cual los parámetros estimados son aquellos que minimizan la suma de las diferencias cuadradas entre las observaciones predichas del modelo y los parámetros, y los datos observados.

Una alternativa a los mínimos cuadrados es la máxima verosimilitud. La máxima verosimilitud elige los valores de los parámetros que maximizan la probabilidad que las observaciones reales hubieran producido si los parámetros fueran ciertos (Hilborn & Walters, 1992).

La elección del método de estimación es importante pues cumple el rol de función objetivo a ser optimizado.

El presente estudio se centra en el método de mínimos cuadrados y el método de máxima verosimilitud.

#### 3.4.1. Método de mínimos cuadrados

El método de mínimos cuadrados es el método de estimación más antiguo y más ampliamente utilizado. Debido, en parte, al hecho que puede ser aplicada directamente de una manera *ad hoc* a los modelos determinísticos, sin ningún tipo de conocimiento siendo tomado de la distribución de probabilidad de las observaciones. (Bard, 1974)

Los mínimos cuadrados son populares para plantear problemas inversos, ya que conducen a cálculos más fáciles. El único inconveniente es su falta de robustez, es decir, su gran sensibilidad a un pequeño número de errores grandes (outliers) en un conjunto de datos (Tarantola, 2005).

**Definición 3.4.1.**

$$F(x) = \frac{1}{2} \sum_{i=1}^m (f_i(x))^2, \quad (3.55)$$

donde  $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $i = 1, \dots, m$  son funciones dadas, y  $m \geq n$ .

Partiendo de la hipótesis inicial que todas las incertidumbres en el problema pueden ser modeladas usando una distribución Gaussiana, cuando la ecuación que resuelve el problema directo es lineal, la expresión explícita obtenida para la distribución de probabilidad también es Gaussiana. Cuando la ecuación es no lineal, la posterior probabilidad no es Gaussiana, pero si la no linealidad no es tan severa, encontrando el punto de máxima verosimilitud de la distribución y estimando la forma de la distribución alrededor de ese punto (i.e., estimando la matriz de covarianza de la distribución), se puede resolver satisfactoriamente el problema (Tarantola, 2005).

Gracias a la simplificación de este resultado, podemos olvidar que existen dos fuentes diferentes de incertidumbre en el espacio de los datos.

El Teorema de Gauss-Markov demuestra que, para problemas lineales, el estimado de mínimos cuadrados tiene la varianza mínima entre todos los estimadores que son funciones lineales de los datos observados ( $d_{obs}$  y  $m_{prior}$ ), independientemente de la forma particular de las funciones de densidad de probabilidad de los datos predichos (d y m-modelo).

Esto no es tan bueno como parece: una mínima varianza puede ser una mala elección cuando las densidades de probabilidad están lejos de ser Gaussianas, como por ejemplo, cuando un pequeño número de outliers no controlados están presentes.

Las teorías físicas (y biológicas) nos permiten hacer predicciones: dada una descripción completa de un sistema, podemos predecir el resultado de algunas medidas.

El problema de predecir el resultado de medidas es llamado el problema de modelización, problema de simulación, o problema directo. El problema inverso consiste en usar el resultado actual de alguna medida para inferir los valores de los parámetros que caracterizan el sistema.

Mientras que el problema directo en física determinística tiene solución única, el problema inverso no. Por esta razón, en los problemas inversos, uno necesita hacer explícita alguna información *a priori* disponible. También se necesita ser cuidadoso en la representación de las incertidumbres de los datos.

### 3.4.2. Método de máxima verosimilitud

Si tenemos los valores observados  $O_i$ ,  $i = 1, \dots, n$ , y queremos estimar el parámetro desconocido  $\theta$  que ajusta un experimento o fenómeno en un modelo, se asume que los valores predichos  $P_i$ ,  $i = 1, \dots, n$  siguen una distribución determinada pues un modelo puede explicarse como la colección de funciones de densidad  $f(O, \theta)$ .

La verosimilitud es definida como la función de densidad evaluada en los datos observados, considerandola únicamente como función del parámetro desconocido  $\theta$  (Millar, 2011).

La verosimilitud basada en distribución normal, asumiendo independencia, es el producto de la verosimilitud en cada dato (o punto).

Para estimar los parámetros del modelo, se maximiza la función de verosimilitud.

$$L(\theta|\text{data}) = \prod_i \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(O_i - P_i)^2}{2\sigma^2}\right] \quad (3.56)$$

Por comodidad, se trabaja con el negativo del logaritmo.

Para estimar los valores de los parámetros, se minimiza la función de log-verosimilitud negativa

$$-\ln L(\theta|\text{data}) = \sum_i \left[ \ln(\sigma) + 0,5 \ln(2\pi) + \frac{(O_i - P_i)^2}{2\sigma^2} \right] \quad (3.57)$$

El término  $0,5 \ln(2\pi)$  es una constante aditiva independiente de los parámetros, por lo que se puede obviar en la ecuación para efectos de la minimización (optimización), (Deriso et al., 2007).

## 3.5. Método de Montecarlo

El método de Montecarlo es un método numérico que permite resolver problemas matemáticos mediante la simulación de variables aleatorias, (Sóbol, 1976). Se realiza una prueba aleatoria, después se repite  $N$  veces de modo que cada experimento sea independiente de los restantes y se toma la media de los resultados de todos los experimentos. El error es proporcional a la magnitud  $\sqrt{\frac{D}{N}}$ , donde  $D$  es una constante y  $N$  es el número de pruebas. (Poca exactitud 5 – 10 %). No se conoce el valor que toma la variable en

un caso concreto dado, pero se sabe qué valores puede tomar y también la probabilidad de los datos. En base a ello, no se puede predecir con exactitud el resultado de una prueba relacionada con esta variable aleatoria, pero si se puede prever con gran seguridad los resultados de un gran número de pruebas. Cuanto mayor sea el número de pruebas, mayor exactitud tendrán las predicciones.

La formulación probabilística de problemas inversos nos lleva a la definición de una distribución de probabilidad en el espacio modelo. (Mosegaard & Tarantola, 1995). Esta distribución de probabilidad combina una información previa con información nueva obtenida mediante la medición de algunos parámetros observados (datos). Como en el caso general, la teoría que conecta los datos con los parámetros del modelo es no lineal, la probabilidad a posteriori en el espacio modelo puede no ser fácil de describir (puede ser multimodal, algunos momentos pueden no ser definidos, etc.).

### 3.6. Algunos modelos de dinámica de poblaciones de peces

En dinámica de poblaciones, los modelos matemáticos permiten la estimación de parámetros. Los problemas en dinámica de poblaciones cuentan con una variedad de modelos.

En el IMARPE, los principales modelos son los clásicos, tales como el modelo Longitud–Peso, el modelo de crecimiento de von Bertalanffy y el modelo de reclutamiento de Ricker.

#### 3.6.1. Modelo Longitud–Peso

El crecimiento es un proceso tridimensional, con longitud, anchura y altura cambiando a la vez.

**El modelo de crecimiento isométrico** La longitud, anchura y altura de un individuo cambia en proporción al otro. Siendo  $L$  la longitud del individuo (u otra medida unidimensional) y  $W$  el peso, tenemos

$$W = aL^3 \tag{3.58}$$

donde  $a$  es un parámetro que se determina a partir de los datos.

En la pesquería, en muchos casos la relación cúbica no ajusta los datos. Una generalización para este modelo es el modelo de crecimiento alométrico, el cual asume que el crecimiento de una dimensión es proporcional a una potencia de otro.

$$W = aL^b \quad (3.59)$$

Mínimos cuadrados para determinar  $a$  y  $b$  en:

$$W = aL^b \quad (3.60)$$

donde

$W$  = peso del ejemplar en gramos

$L$  = longitud del ejemplar en centímetros

$a$  = factor de condición (coeficiente de variación de Fulton)

$b$  = constante de crecimiento

Linealizando la ecuación (3.60) tenemos:

$$\ln W = \ln a + b \ln L \quad (3.61)$$

de ahí, por la Proposición 3.6.1 (abajo), obtenemos:

$$b = \frac{n \sum (\ln L \ln W) - \sum (\ln L) \sum (\ln W)}{n \sum (\ln L)^2 - \left( \sum \ln L \right)^2} \quad (3.62)$$

$$\ln a = \overline{\ln W} - b \overline{\ln L} \quad (3.63)$$

donde

$n$  = número de datos

**Proposición 3.6.1.** (Simplificación del modelo) Sean los puntos  $(x_i, y_i) \in \mathbb{R}^2$  con  $i = 1, \dots, n$ ,  $n$  entero. Queremos ajustar dichos puntos a una recta  $y = a + bx$ . Luego, aplicando el método de mínimos cuadrados tenemos:

$$b = \frac{n \sum_{i=1}^n (x_i y_i) - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2} \quad (3.64)$$

$$a = \bar{y}_i - b\bar{x}_i \quad (3.65)$$

*Demostración.* Sea la función  $G(a, b) = \sum_{i=1}^n (y_i - a - bx_i)^2$ .

Luego, por el método de mínimos cuadrados queremos minimizar  $G(a, b)$ .

Derivando e igualando a 0 obtenemos:

$$\frac{dG}{da} = 2 \sum_{i=1}^n (y_i - a - bx_i)(-1) = 0$$

$$\frac{dG}{db} = 2 \sum_{i=1}^n (y_i - a - bx_i)(-x_i) = 0$$

de ahí, como

$$-2 \sum_{i=1}^n (y_i - a - bx_i) = 0$$

y

$$-2 \sum_{i=1}^n (y_i - a - bx_i)(x_i) = 0$$

tenemos

$$\begin{aligned} \sum_{i=1}^n (y_i - a - bx_i) &= \sum_{i=1}^n y_i - \sum_{i=1}^n a - \sum_{i=1}^n bx_i \\ &= \sum_{i=1}^n y_i - na - b \sum_{i=1}^n x_i = 0 \end{aligned} \quad (3.66)$$

y

$$\begin{aligned} \sum_{i=1}^n (y_i - a - bx_i)(x_i) &= \sum_{i=1}^n (x_i y_i - ax_i - bx_i^2) \\ &= \sum_{i=1}^n x_i y_i - a \sum_{i=1}^n x_i - b \sum_{i=1}^n x_i^2 = 0 \end{aligned} \quad (3.67)$$

de (3.66) y (3.67):

$$b \sum_{i=1}^n x_i = \sum_{i=1}^n y_i - na$$

y

$$b \sum_{i=1}^n x_i^2 = \sum_{i=1}^n (x_i y_i) - a \sum_{i=1}^n x_i$$

entonces

$$b \left( \sum_{i=1}^n x_i \right)^2 = \sum_{i=1}^n x_i \sum_{i=1}^n y_i - na \sum_{i=1}^n x_i$$

y

$$bn \sum_{i=1}^n x_i^2 = n \sum_{i=1}^n (x_i y_i) - na \sum_{i=1}^n x_i$$

Combinando las ecuaciones:

$$\begin{aligned} bn \sum_{i=1}^n x_i^2 - b \left( \sum_{i=1}^n x_i \right)^2 &= n \sum_{i=1}^n (x_i y_i) - \sum_{i=1}^n x_i \sum_{i=1}^n y_i \\ b \left( n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2 \right) &= n \sum_{i=1}^n (x_i y_i) - \sum_{i=1}^n x_i \sum_{i=1}^n y_i \\ b &= \frac{n \sum_{i=1}^n (x_i y_i) - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2} \end{aligned} \quad (3.68)$$

De (3.66):

$$\begin{aligned} na &= \sum_{i=1}^n y_i - b \sum_{i=1}^n x_i \\ a &= \frac{1}{n} \left( \sum_{i=1}^n y_i - b \sum_{i=1}^n x_i \right) = \frac{1}{n} \sum_{i=1}^n y_i - \frac{b}{n} \sum_{i=1}^n x_i \\ \Rightarrow a &= \bar{y} - b\bar{x} \end{aligned} \quad (3.69)$$

□

### 3.6.2. Modelo de crecimiento o de Longitud-Edad

Este modelo, propuesto por von Bertalanffy, es el más estudiado y más usado de los modelos de crecimiento. Llamado así en honor a Ludwig von Bertalanffy, quien desarrollo este modelo para el proceso de crecimiento humano. Esta clase de modelo es la más estudiada y la más usada de todos los modelos de peces por su utilidad y porque describe de forma satisfactoria el crecimiento. El principio fundamental de su éxito es que la variación en

el crecimiento de los peces tiende a ser más lento con la edad. Este modelo describe la relación de longitud ( $L$ ) con la edad ( $t$ ), aunque también es apropiado para otras variables de tamaño, como por ejemplo, peso ( $W$ ).

El modelo más simple puede ser escrito para el caso de disminución de crecimiento respecto a la edad como una ecuación diferencial con una disminución lineal de la variación de crecimiento.

$$\frac{dL}{dt} = \omega - kL, \quad \omega > 0, k > 0 \quad (3.70)$$

con la condición inicial  $L(t_0) = L_0$  (Quinn & Deriso, 1999; Gallucci & Quinn, 1979).

El parámetro  $\omega$  es la variación de crecimiento en el tiempo inicial  $t_0$ . El parámetro  $k$  es un parámetro de crecimiento, con unidades de  $t^{-1}$ .

La ecuación diferencial de crecimiento especifica una variación de crecimiento inicial  $\omega$ , el cual decrece linealmente como una función de longitud hacia 0 desde una longitud  $\frac{\omega}{k} = L_\infty$ , el cual es la longitud asintótica.

Reemplazando  $\frac{\omega}{k} = L_\infty$  en (3.70):

$$\frac{dL}{dt} = k(L_\infty - L), \quad k > 0, \quad L_\infty > 0 \quad (3.71)$$

con la condición inicial  $L(t_0) = L_0$ .

La solución a la ecuación diferencial (Gallucci & Quinn, 1979):

$$\begin{aligned}
 \frac{dL}{dt} &= k(L_\infty - L) \\
 \frac{dL}{L_\infty - L} &= k dt \\
 \int_{L_0}^L \frac{dL}{L_\infty - L} &= \int_{t_0}^t k dt \\
 \ln(L_\infty - L_0) - \ln(L_\infty - L) &= k(t - t_0) \\
 \ln\left(\frac{L_\infty - L_0}{L_\infty - L}\right) &= k(t - t_0) \\
 \frac{L_\infty - L_0}{L_\infty - L} &= e^{k(t-t_0)} \\
 L_\infty - L_0 &= L_\infty e^{k(t-t_0)} - L e^{k(t-t_0)} \\
 L e^{k(t-t_0)} &= L_0 + L_\infty [e^{k(t-t_0)} - 1] \\
 L &= L_0 e^{-k(t-t_0)} + L_\infty [1 - e^{-k(t-t_0)}] \\
 L(t) &= L_\infty [1 - e^{-k(t-t_0)}] + L_0 e^{-k(t-t_0)} \quad (3.72)
 \end{aligned}$$

Al evaluar la ecuación (3.72) para  $t = t_0$  obtenemos  $L(t_0) = L_0$ . Usualmente se elegiría  $t_0 = 0$  con  $L_0 > 0$ , pero para simplificar la expresión, hacemos uso de un arreglo con el cual elegimos  $L_0 = 0$  e interpretamos  $t_0$  como la edad del individuo cuando este tenía "longitud 0". Obteniendo la expresión:

$$L(t) = L_\infty [1 - e^{-k(t-t_0)}] \quad (3.73)$$

llamado modelo de crecimiento von Bertalanffy en honor a Ludwig von Bertalanffy, quien desarrollo este modelo para el proceso de crecimiento humano.

### 3.6.3. Modelo Stock-Reclutamiento de Ricker

Este modelo es principalmente usado cuando el canibalismo de jóvenes por adultos es un importante mecanismo de regulación, o cuando mayores densidades aumentan el tiempo necesario por los peces jóvenes para crecer a través de un rango de tamaños particularmente vulnerable, o cuando existe un desfase en la respuesta de un depredador o un parásito a la abundancia de los peces jóvenes que éste consume, con la consiguiente sobrecompensación de densidades iniciales más altas de las especies presa (Ricker, 1975).

Para alguna población de peces, el ciclo de regeneración puede ser visualizado como

Huevos  $\rightarrow$  Larvas  $\rightarrow$  Juveniles  $\rightarrow$  Reclutas  $\rightarrow$  Desovantes  $\rightarrow$  Huevos  $\rightarrow \dots$

La abundancia en cada estadio se asume proporcional al estadio previo. Así, la producción de huevos  $N_0$  es proporcional al stock desovante  $S$ ,

$$N_0 = fS, \quad (3.74)$$

donde  $f$  es la fecundidad neta promedio de la población. El reclutamiento  $R$ , es proporcional a la producción de huevos

$$R = lN_0, \quad (3.75)$$

donde  $l$  es la supervivencia de vida temprana desde el estadio de huevo hasta el momento del reclutamiento. De ahí, el reclutamiento es también proporcional al stock desovante.

$$R = \alpha S, \quad (3.76)$$

donde  $\alpha = lf$ . El parámetro  $\alpha$  es llamado parámetro de productividad o denso-independiente.

La relación (3.76), en general, no es realística porque esto nos indica que el reclutamiento aumenta sin límite como una función del stock desovante, a menos que la fecundidad sea denso-dependiente.

Si los efectos denso-dependientes están presentes en algún estadio de vida, y considerando  $Z_t$  como la mortalidad en el estadio de vida  $t$  y  $T$  la edad del reclutamiento, la función del tamaño de la población puede definirse como

$$\frac{dN}{dt} = -Z_t N, \quad \text{donde } N(0) = N_0 \quad (3.77)$$

donde  $Z_t$  está compuesto de efectos denso-dependientes y denso-independientes, lo cual puede ser representado por

$$Z_t = a + bS \quad (3.78)$$

Luego, la ecuación (3.77) es de la forma

$$\frac{dN}{dt} = -(a + bS)N \quad (3.79)$$

cuya solución es:

$$N(t) = N_0 e^{-(a+bS)t} \quad (3.80)$$

Para la edad del reclutamiento,  $t = T$ , y por la ecuación (3.74) tenemos

$$\begin{aligned} R = N(T) &= f S e^{-aT} \cdot e^{-bST} \\ &= (f e^{-aT}) S e^{-bTS} \\ &= \alpha S e^{-\beta S}, \end{aligned} \quad (3.81)$$

el cual es llamado la curva stock–reclutamiento de Ricker.

El modelo de Ricker relaciona el stock y reclutamiento de la forma:

$$R = \alpha S S B e^{-\beta S S B} \quad (3.82)$$

### 3.6.4. Modelo estadístico de evaluación del stock basado en captura por edad y por tallas

Los modelos de evaluación se pueden clasificar según los datos ingresados al modelo, cada uno de los cuales tienen ventajas y desventajas, (ICES, 2012; Maunder & Punt, 2013).

Modelos de evaluación basados en capturas por edad reconstruyen la estructura de la población a través de estos datos, reconstruyendo una serie de tiempo de la biomasa y biomasa desovante que debió tener la población para que haya sido posible los niveles de captura ingresados como dato en el modelo. En estos modelos, también se estima información adicional, tales como la tasa instantánea de mortalidad por pesca el número de reclutas.

Para estimar la abundancia ( $N$ ), biomasa ( $B$ ), biomasa desovante ( $BD$ ) y reclutamientos ( $R$ ), así como la tasa instantánea de mortalidad por pesca ( $F$ ) anual del stock de una especie para un periodo dado, se procede de la siguiente forma (Deriso, 1980):

La abundancia se estima para todos los años y todas las edades siguiendo el decaimiento exponencial debido a la mortalidad natural y mortalidad por pesca:

$$N_{y+1,a+1} = N_{y,a} e^{-(M+F_{y,a})} \quad (3.83)$$

donde  $N_{y,a}$  es la abundancia en el año  $y$  de la edad  $a$ ,  $M$  es la mortalidad natural que fue asumida constante para todos los años y todas las edades y  $F_{y,a}$  es la tasa instantánea de mortalidad por pesca.

La biomasa total anual se estima como la sumatoria de los productos de las abundancias  $N_{y,a}$  por el vector de pesos a la edad  $W_a$ :

$$B_y = \sum_{a=1}^{12} (N_{y,a} W_a) \quad (3.84)$$

donde  $W_a$  se considera constante para todos los años.

La biomasa desovante se estima como la sumatoria de los productos de las abundancias, el vector de pesos a la edad, la ojiva de madurez  $MS_a$  y la fracción del año donde ocurre el desove:

$$BD_y = \sum_{a=1}^{12} (N_{y,a} W_a MS_a e^{-\Delta_s Z_{y,a}}) \quad (3.85)$$

donde  $MS_a$  es la ojiva de madurez, considerada constante para toda la serie de tiempo, y  $\Delta_s$  es la fracción del año donde ocurre el desove.

El reclutamiento anual se considerado como la abundancia a edades determinadas, por ejemplo a la edad 1:

$$R_y = N_{y,1} \quad (3.86)$$

Los desembarques anuales se estiman como la sumatoria de la biomasa capturada de todas las edades:

$$C_y = \sum_{a=1}^{12} \left( N_{a,y} W_{a,y} \frac{F_{a,y}}{Z_{a,y}} (1 - e^{-Z_{a,y}}) \right) \quad (3.87)$$

Estos modelos suelen optimizarse al maximizar las verosimilitudes entre la biomasa media y la biomasa estimada por el método acústico, entre los reclutamientos medios y un reclutamiento constante y entre los desembarques predichos y los desembarques observados. Asumiendo que las diferencias entre las estimaciones y las observaciones (errores) siguen una distribución lognormal, la función de verosimilitud por variable se estima como:

$$\ln L(\gamma|\theta) = \frac{1}{2cv^2} \sum (\ln \gamma_y - \ln \hat{\gamma}_y)^2 \quad (3.88)$$

donde  $\gamma_y$  es la variable observada y  $\hat{\gamma}_y$  es la variable predicha por el modelo y  $cv$  es el coeficiente de variación o peso de cada variable dentro del proceso de optimización.

# Capítulo 4

## Materiales y Métodos

El estudio se enfocó en los modelos clásicos empleados en dinámica de poblaciones: el modelo longitud–peso, el modelo longitud–edad y el modelo stock–reclutamiento.

El estudio se realizó en tres etapas: La primera etapa fue de preparación y desarrollo, donde se definieron los límites del estudio y se obtuvieron datos que posteriormente fueron utilizados en la siguiente etapa. La segunda etapa fue de validación, se trabajó directamente con los datos y se obtuvieron los resultados para poder comparar diferentes métodos de estimación y algoritmos de optimización. Con los resultados obtenidos, en la tercera etapa se realizó una aplicación.

A continuación se detalla los materiales utilizados, así como la metodología usada.

### 4.1. Datos

Para la realización del trabajo se consideraron dos tipos de datos. En la primera etapa se utilizaron datos simulados de cada parámetro en los modelos de longitud–talla, longitud–edad y stock–reclutamiento. En la segunda y tercera etapa se emplearon datos reales proporcionados por el *Instituto del Mar del Perú* (IMARPE).

### 4.1.1. Datos simulados

Se realizó un experimento previo para determinar el número adecuado de repeticiones que se debería utilizar en las simulaciones para garantizar la representatividad del set de datos a generar.

Para la generación de datos simulados se utilizó el método de Montecarlo.

Teniendo como base valores aproximados de los parámetros de crecimiento de anchoveta, los fijamos para utilizarlos como valores semilla. A cada uno de estos parámetros se le agregaron distintos grados de error, asegurándonos que estos tengan sentido biológico.

El tipo de error introducido en la generación de los datos simulados se basan en los tipos de error más frecuentes ocurridos en la toma de datos. Para ello se procedió a introducir dos tipos de error:

- Introduciendo error de observación (ruido lognormal multiplicativo)
- Introduciendo error de procesos (ruido en los parámetros)

Se consideró adecuado controlar el error a través de la variación de la desviación estándar de la distribución log-normal que se le multiplicó al valor semilla.

Se realizaron simulaciones hasta el momento en que se estabilizó el error relativo de la media y la desviación estándar respecto a los parámetros iniciales.

Se observó que la estabilización se aseguraba al llegar a utilizar 1000 valores simulados, por lo que a partir de este momento se trabajaron con sets de datos con 1000 valores.

### 4.1.2. Datos observados

Para el modelo longitud–peso se contó con datos biométricos y parámetros biológicos de anchoveta provenientes del crucero de investigación de recursos pelágicos realizado por el IMARPE entre noviembre y diciembre del 2010 (Cr1011-12).

Para el modelo longitud–edad, se utilizaron los parámetros de crecimiento de anchoveta estimados por Pauly & Tsukayama (1983), y los cuales se citan en Pauly & Tsukayama (1987) y Pauly et al. (1989).

Para el modelo estadístico de captura por edad y captura por talla, se contó con la serie histórica de los desembarques anuales registrados de jurel desde 1970 hasta 2012, estimaciones de biomasa por métodos acústicos desde 1970 hasta 2012 y estimaciones de captura por unidad de esfuerzo (CPUE), además de la frecuencia de tallas de jurel obtenidos del seguimiento de la pesquería.

## 4.2. Software

Para estudiar cada modelo se empleó un software y herramienta específica, dentro de los cuales están implementados cada uno de los algoritmos de optimización que se utilizaron en el estudio.

El estudio consideró los tres softwares más utilizados por el IMARPE en la estimación de parámetros de los modelos de crecimiento, las herramientas en cada software se especifican a continuación:

- **Solver–Excel:** Se utilizó el *MS Excel* con la herramienta *solver*.  
En dicha herramienta está implementado el método *Gradiente Conjugado*, el cual utiliza un cálculo numérico para estimar las derivadas.  
Este software es el más utilizado en el Instituto del Mar del Perú.
- **Optim–R:** Se utilizó el software estadístico *R* (R Development Core Team, 2012) con la función *optim*.  
En dicha función está implementado los métodos *Quasi-Newton*, *BFGS* y *L-BFGS-B*, los cuales utilizan un cálculo numérico para estimar las derivadas. También está implementado el método *Nelder-Mead*, el cual es un método de optimización determinista libre de derivadas.
- **AD Model Builder:** Se utilizó el *AD Model Builder* (Bolker et al., 2013).  
En dicha herramienta está implementado el método *Diferenciación Automática*, el cual utiliza un cálculo de derivadas exacta, que recoge resultados intermedios y realiza cálculos internos basados en la regla de la cadena.

## **4.3. Modelos matemáticos en dinámica poblacional**

El presente trabajo se enfocó en los modelos clásicos empleados en dinámica de poblaciones: el modelo longitud–peso, el modelo longitud–edad y el modelo stock–reclutamiento.

### **4.3.1. Modelo Longitud–Peso**

Para este modelo, se utilizaron los métodos de estimación de mínimos cuadrados y máxima verosimilitud, y los métodos de optimización Gradiente Conjugado, Nelder–Mead, BFGS, L–BFGS–B, Diferenciación Automática.

Este caso es especial pues, asumiendo error log-normal, es el único que tiene solución analítica.

### **4.3.2. Longitud–Edad (von Bertalanffy)**

Para este modelo, se utilizaron los métodos de estimación de mínimos cuadrados y máxima verosimilitud, y los métodos de optimización Gradiente Conjugado, Nelder–Mead, BFGS, L–BFGS–B, Diferenciación Automática.

### **4.3.3. Stock–Reclutamiento (Ricker)**

Para este modelo, se utilizaron los métodos de estimación de mínimos cuadrados y máxima verosimilitud, y los métodos de optimización Gradiente Conjugado, Nelder–Mead, BFGS, L–BFGS–B, Diferenciación Automática.

## **4.4. Metodología**

Habiendo implementado los modelos, se quería determinar si estos devuelven las soluciones esperadas (valores próximos a las soluciones reales). Para ello, se procedió a realizar una simulación mediante el método de Montecarlo, para lo cual se procedió a generar varios sets de datos aleatorios.

Para determinar el número adecuado de sets de valores (datos) simulados, se realizó un experimento adicional en el cual, con datos al azar, se generaron

nuevos datos hasta que la media y la desviación estándar de los mismos llegaban a un valor estacionario (estado de equilibrio).

Para generar los valores aleatorios, se procedió de dos maneras:

- Introducir error lognormal multiplicativo. Este método simula el error de observación. A los resultados obtenidos con los valores fijos se le agregó el ruido. Para simular valores de muestreo se agregó un error log-normal multiplicativo. Se eligió utilizar un error log-normal multiplicativo, pues un error normal aditivo no garantiza la obtención de valores con sentido biológico.
- El otro método utilizado para generar los datos simulados es aplicar ruido a cada parámetro (error de procesos).

Partiendo del conjunto de datos simulados por el método de Montecarlo, se procedió a realizar un remuestreo de 1000 repeticiones con distintos grados de error log-normal (variaciones de la desviación estándar desde 0.01 hasta 0.2 con pasos de 0.01).

Se implementaron los modelos con las herramientas anteriormente mencionadas de modo que puedan ser utilizados con los métodos de mínimos cuadrados y máxima verosimilitud.

Con el conjunto de datos obtenidos se estimaron los parámetros de cada repetición. Las medias de los parámetros estimados se compararon con los valores iniciales (validación de los modelos).

Para verificar que el resultado obtenido con el modelo no dependía de los parámetros iniciales, se procedió a comparar los resultados para un mismo parámetro de entrada.

Adicionalmente, para el caso del modelo de longitud–peso, se linealizó la ecuación conjuntamente con los datos a los que se le aplicó el modelo, se evaluaron estos últimos datos con la fórmula de mínimos cuadrados para el caso de una recta, y se compararon los resultados obtenidos (hay que notar que en este caso se asume implícitamente un error log-normal).

## 4.5. Comparación de métodos de estimación

Los métodos de estimación utilizados para cada modelo se explican a continuación:

Modelo	Método de estimación	
	Mínimos cuadrados	Máxima verosimilitud
Longitud-peso	X	X
Longitud-edad	X	X
Stock-reclutamiento	X	X

Cuadro 4.1: Métodos de estimación utilizados para cada modelo.

## 4.6. Comparación de algoritmos de optimización

Para comparar los métodos de optimización, se analizó la estabilidad del algoritmo recuperando los parámetros ajustando los datos sin error. Esto sirve de verificación para (i) la implementación del código de la función objetivo y (ii) el algoritmo de optimización.

Método	Derivada	Software
Gradiente Conjugado	Numérica	Solver-Excel
Nelder-Mead	Numérica	Optim-R
BFGS	Numérica	Optim-R
L-BFGS-B	Numérica	Optim-R
Diferenciación automática	Exacta	ADMB

Cuadro 4.2: Tipo de derivada para cada método de optimización y software donde se encuentra implementado.

## 4.7. Aplicación de métodos de estimación y optimización a la calibración de un modelo de dinámica poblacional del jurel (*Trachurus murphyi*) frente a Perú

Una aplicación directa de este trabajo es el modelo aplicado a la evaluación del recurso del jurel, en el marco de la reunión de la *Organización Regional de Ordenación Pesquera del Pacífico Sur* (OROP-PS) (SPRFMO, 2012,0; Ramos-Vásquez, 2013).

El modelo ajusta varios aspectos de evaluación tales como biomasa total, utiliza forzantes, tales como series de tiempo de capturas, e índices, tales como biomasa estimada por métodos acústicos y captura por unidad de esfuerzo. Además, incorpora información de captura anual y estructura de tamaños.

Este modelo está implementado en *AD Model Builder* y utiliza el método de optimización de *diferenciación automática* y el método de estimación de *máxima verosimilitud*.

La función objetivo del modelo es una función de verosimilitud que ajusta diversos parámetros estimados de la biología del recurso, como son los índices de abundancia, reclutamiento y selectividad, considerando el ajuste de los desembarques.

Componentes de la función de verosimilitud:

- Índice de abundancias:

$$L_1 = \sum_s \lambda_1^s \sum_i \log \left( \frac{I_i^s}{\hat{I}_i^s} \right)^2 \quad (4.1)$$

- Prior de suavizamiento de las selectividades:

$$L_2 = \sum_l \lambda_2^l \sum_{j=1}^{12} (\eta_{j+2}^l + \eta_j^l - 2\eta_{j+1}^l)^2 \quad (4.2)$$

- Prior de regularidad del reclutamiento:

$$L_3 = \lambda_3 \sum_{i=1958}^{2012} \varepsilon_i^2 \quad (4.3)$$

- Función de verosimilitud de biomasa capturada:

$$L_4 = \sum_f \lambda_4^f \sum_{i=1970}^{2012} \log \left( \frac{C_i^f}{\hat{C}_i^f} \right)^2 \quad (4.4)$$

- Función de verosimilitud de proporción a la edad/longitud:

$$L_5 = - \sum_{\nu, i, j} n^\nu P_{i,j/l}^\nu \log(\hat{P}_{i,j/l}^\nu) \quad (4.5)$$

- Ajuste de la curva de reclutamiento:

$$L_6 = \lambda_6 \sum_{i=1970}^{2012} \log \left( \frac{N_{i,1}}{\hat{R}_i} \right)^2 \quad (4.6)$$

Función objetivo a ser minimizada:

$$\dot{L} = \sum_k L_k \quad (4.7)$$

# Capítulo 5

## Resultados y Discusión

### 5.1. Resultados

Los resultados obtenidos se detallan a continuación:

#### 5.1.1. Simulación de Montecarlo

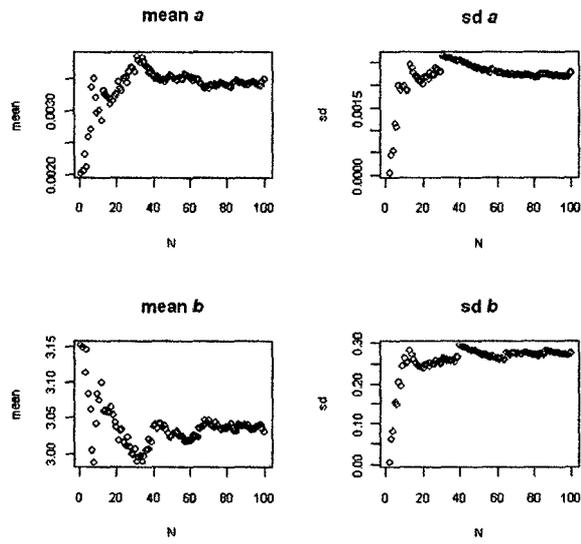
En el análisis exploratorio con simulaciones Montecarlo que se efectuó para los modelos Longitud–Peso y Longitud–Edad arrojó como resultado que con un número de 1000 iteraciones se obtendría un ajuste razonable al valor esperado, pues en valores cercanos a 1000 se alcanza estabilidad en el proceso de estimación de la media, además de la desviación estándar (Figura 5.1 y Figura 5.2).

En la Figura 5.1(a) aún se observa variabilidad en los estimados de  $a$  y  $b$ , mientras que en la Figura 5.1(b) ya se observa estacionalidad.

En el caso de las desviaciones estándar sucede lo mismo, por lo que es recomendable utilizar 1000 como número de repeticiones.

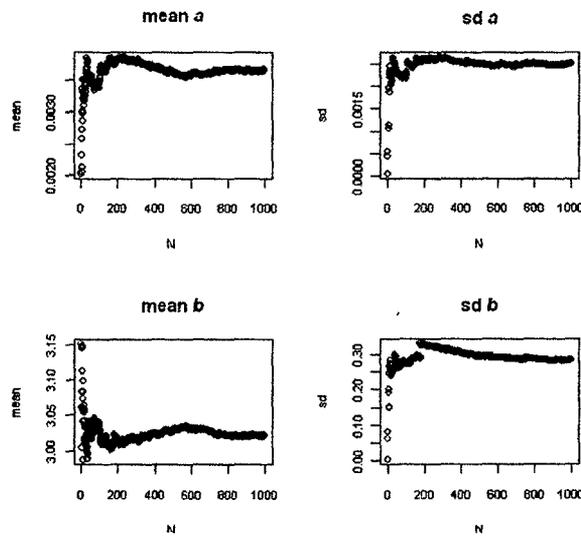
En el caso de la Figura 5.2 sucede lo mismo. En la Figura 5.2(a) se observa una variabilidad significativa en los estimados de la media y desviaciones estándar de  $L_{\infty}$ ,  $K$  y  $t_0$ , por lo que se decidió utilizar 1000 iteraciones ya que a ese número se alcanza estabilidad de los estimados (Figura 5.2(b)).

**Longitud-Peso**



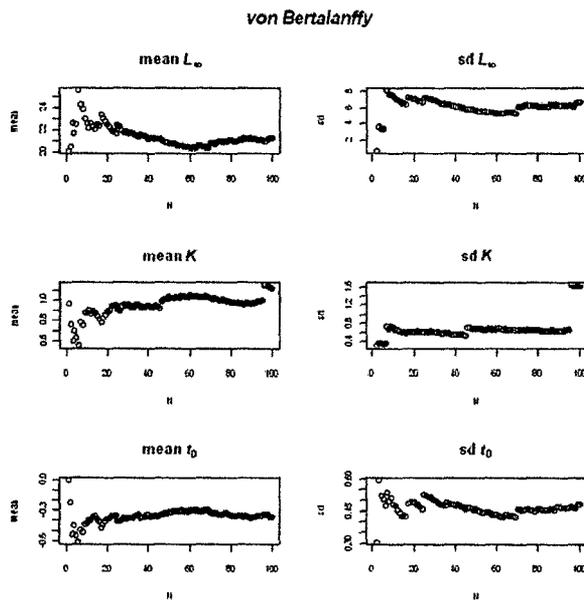
(a) Resultados de las simulaciones Montecarlo con 100 repeticiones (n=100)

**Longitud-Peso**

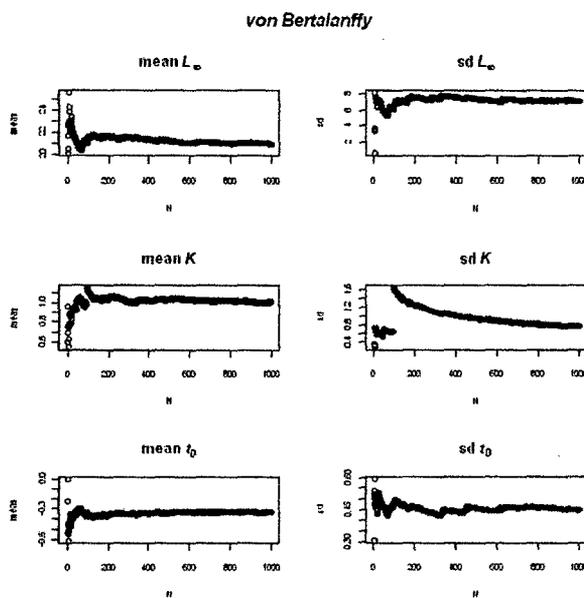


(b) Resultados de las simulaciones Montecarlo con 1000 repeticiones (n=1000)

Figura 5.1: Simulaciones Montecarlo de los parámetros para el modelo Longitud-Peso.



(a) Resultados de las simulaciones Montecarlo con 100 repeticiones ( $n=100$ )



(b) Resultados de las simulaciones Montecarlo con 1000 repeticiones ( $n=1000$ )

Figura 5.2: Simulaciones Montecarlo de los parámetros para el modelo von Bertalanffy.

### 5.1.2. Longitud–peso

Conocido los resultados anteriores, para estimar los parámetros  $a$  y  $b$  del modelo Longitud-Peso se realizaron las 1000 simulaciones para cada grado de error (desviaciones estándar de 0.01 hasta 0.20, con pasos de 0.01), tomando como base para los valores de los parámetros obtenidos del crucero de evaluación de recursos pelágicos realizado entre noviembre y diciembre del 2010.

En la Figura 5.3 se muestra el comportamiento de las medias y desviaciones estándar de los valores obtenidos. Todos los métodos siguen la misma tendencia, a excepción del método de *gradiente conjugado*, el cual presenta menor grado de exactitud respecto a los resultados obtenidos con los métodos *Nelder–Mead*, *BFGS*, *L–BFGS–B* y *diferenciación automática*.

Adicionalmente, para este método se estimó la solución analítica del modelo, Figura 5.4. Aquí observamos que los valores estimados siguen una misma tendencia, sin alejarse demasiado del valor óptimo mientras se aumenta el grado de incertidumbre o ruido en los valores semilla.

En la Tabla 5.1 se resume los resultados obtenidos para cada método de optimización y de estimación. En el caso del método de *Gradiente Conjugado*, se consiguió llegar a una solución después de un número elevado de iteraciones con el método de *mínimos cuadrados*, sin embargo, el método de *máxima verosimilitud* no se pudo implementar en el software debido a su complejidad.

### 5.1.3. Longitud–edad

Se realizaron 1000 simulaciones para cada grado de error (desviaciones estándar variando desde 0.01 hasta 0.20, con pasos de 0.01) para poder estimar los parámetros  $L_\infty$ ,  $K$  y  $t_0$ , tomando como base mediciones biométricas de anchoveta.

En la Figura 5.5 se muestra el comportamiento de las medias y desviaciones estándar de los valores obtenidos. Todos los métodos siguen la misma tendencia. El método *L–BFGS–B* presentó error en la ejecución, por lo que no fue posible utilizarlo en el análisis. No se implementó el método de *máxima verosimilitud* con el método de *Gradiente Conjugado* debido a la complejidad de implementación.

En la Tabla 5.2 se muestran los resultados, se pudo apreciar que las

estimaciones obtenidas no se alejan entre sí, sin embargo, las estimaciones del método *Nelder-Mead* presentaban ligeramente mayor error que los otros métodos.

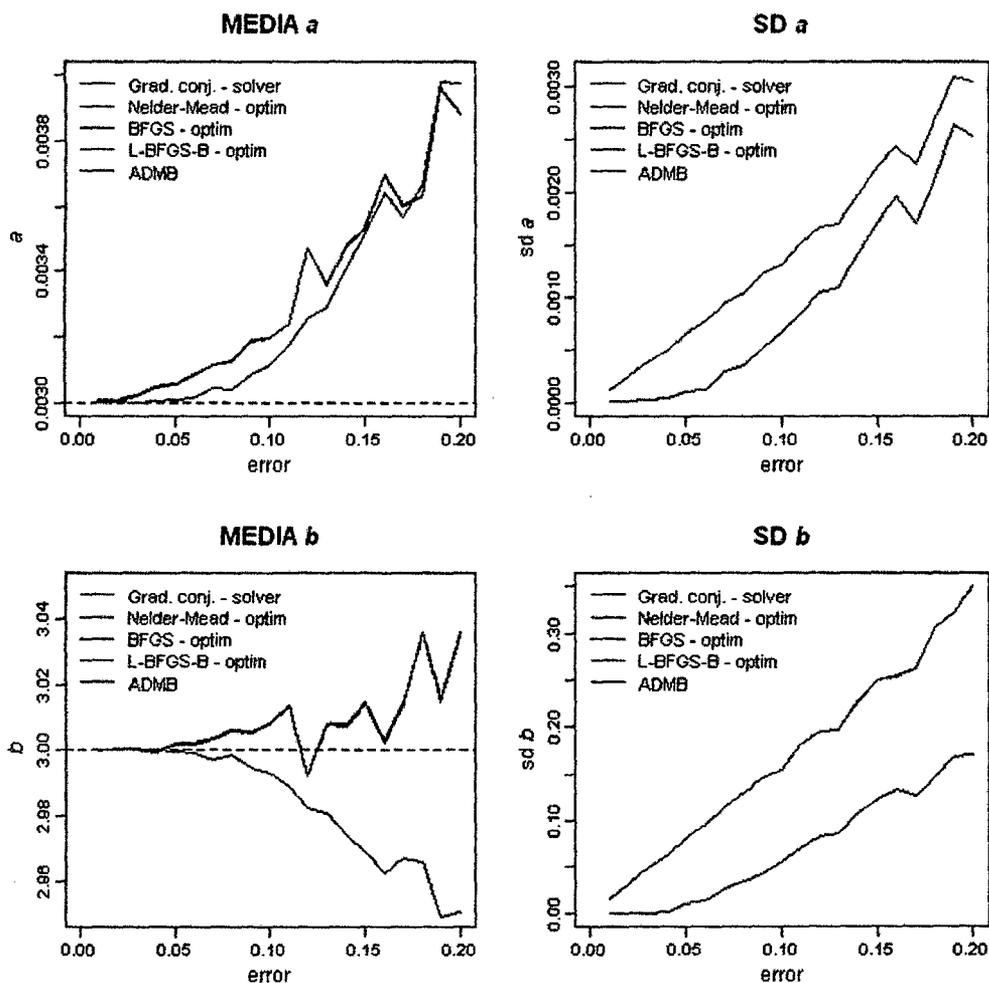


Figura 5.3: Resultados de las estimaciones de los parámetros del modelo Longitud-Peso, para datos con diferente grado de dispersión. La línea discontinua roja indica el valor real del parámetro.

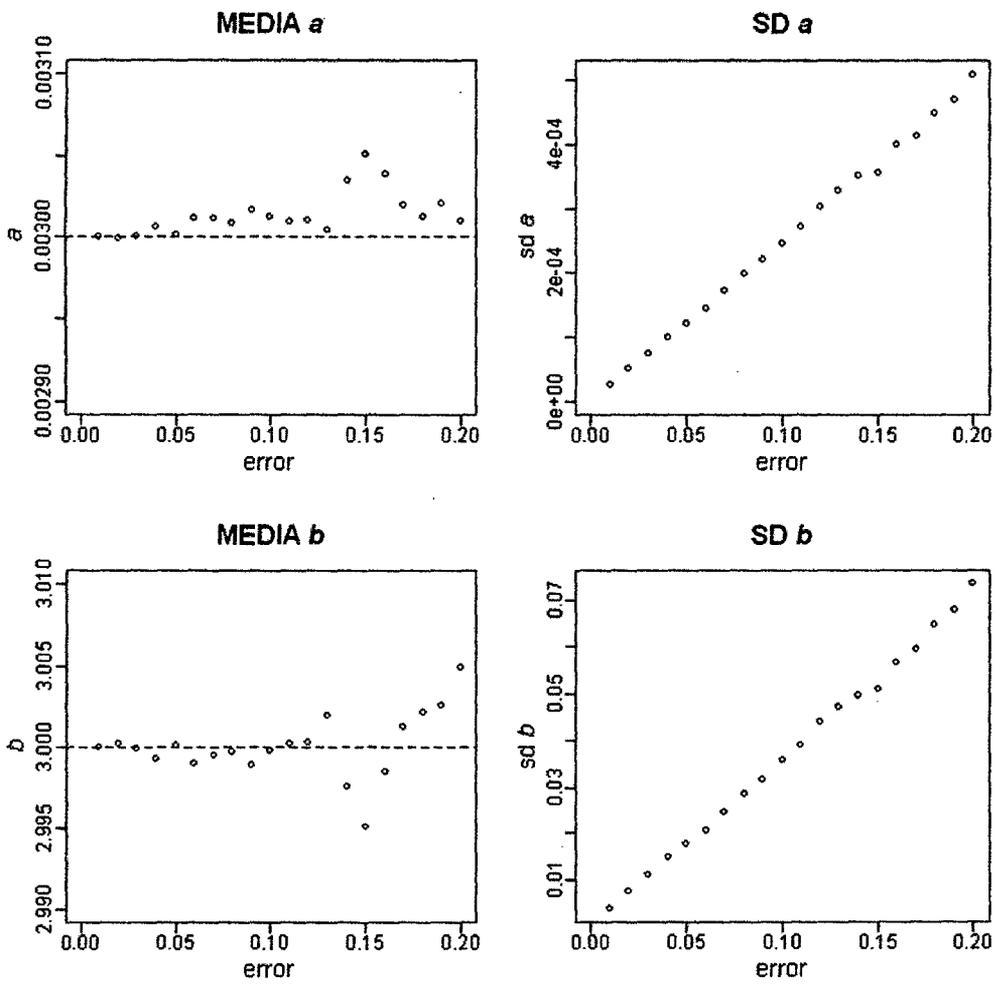


Figura 5.4: Resultados de la solución analítica del modelo Longitud-Peso.

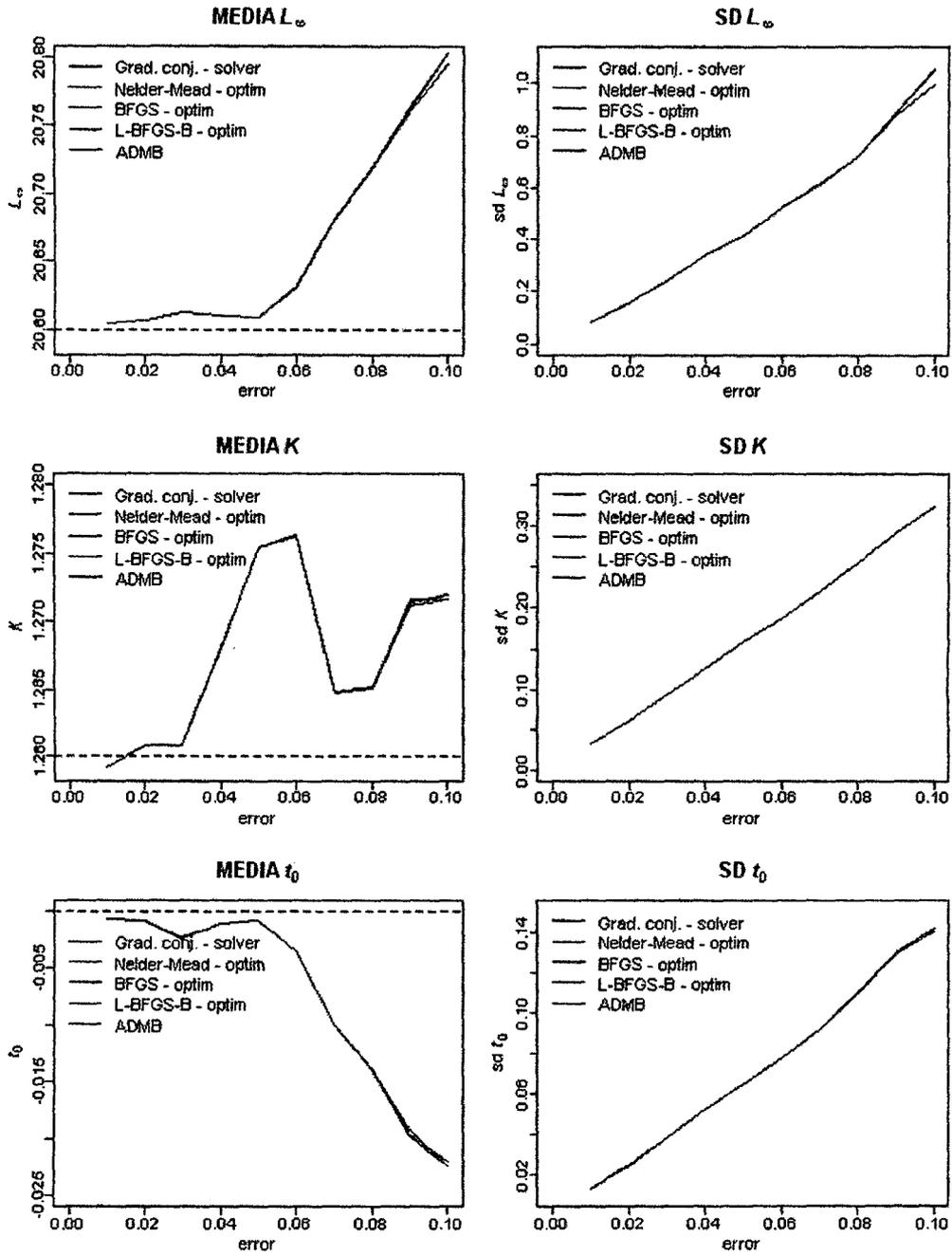


Figura 5.5: Resultados de las estimaciones de los parámetros del modelo Von Bertalanffy, para datos con diferente grado de dispersión. La línea discontinua roja indica el valor real del parámetro.

Método de optimización	Método de estimación	
	mínimos cuadrados	máxima verosimilitud
Gradiente Conjugado (Excel - solver)	número de iteraciones elevado	implementación demasiado compleja
Nelder-Mead (R - optim)	presenta los mejores resultados	presenta los mejores resultados
BFGS (R - optim)	mayor error en la estimación	mayor error en la estimación
L-BFGS-B (R - optim)	mayor error en la estimación	mayor error en la estimación
Diferenciación Automática (AD Model Builder)	presenta los mejores resultados	presenta los mejores resultados

Cuadro 5.1: Cuadro comparativo de los resultados obtenidos por cada método utilizado en el modelo Longitud-Peso.

Método de optimización	Método de estimación	
	mínimos cuadrados	máxima verosimilitud
Gradiente Conjugado (Excel - solver)	comparable con ADMB	implementación demasiado compleja
Nelder-Mead (R - optim)	a mayor error, se aleja más rápido de la solución	a mayor error, se aleja más rápido de la solución
BFGS (R - optim)	mayor error en la estimación	mayor error en la estimación
L-BFGS-B (R - optim)	error al ejecutar	error al ejecutar
Diferenciación Automática (AD Model Builder)	presenta los mejores resultados	presenta los mejores resultados

Cuadro 5.2: Cuadro comparativo de los resultados obtenidos por cada método utilizado en el modelo Von Bertalanffy.

#### 5.1.4. Stock-reclutamiento

Se realizaron 1000 simulaciones para cada grado de error (desviaciones estándar de 0.01 hasta 0.20, con pasos de 0.01) para poder estimar los parámetros  $\alpha$  y  $\beta$  a partir de valores simulados.

En la Figura 5.6 se muestra el comportamiento de las medias y desviaciones estándar de los valores obtenidos. Todos los métodos siguen la misma tendencia, sin embargo, los métodos *BFGS* y *L-BFGS-B* se alejaron más de la solución. El método de *Gradiente Conjugado* presentó mayor variabilidad en los resultados. No se implementó el método de *máxima verosimilitud* con el método de *Gradiente Conjugado* debido a la complejidad de implementación.

En la Tabla 5.3 se muestran los resultados. De aquí, los mejores resultados se obtuvieron con los métodos *Nelder-Mead* y *Diferenciación Automática*.

Método de optimización	Método de estimación	
	mínimos cuadrados	máxima verosimilitud
Gradiente Conjugado (Excel - solver)	presenta mayor variabilidad	implementación demasiado compleja
Nelder-Mead (R - optim)	presenta los mejores resultados	presenta los mejores resultados
BFGS (R - optim)	mayor error en la estimación	mayor error en la estimación
L-BFGS-B (R - optim)	mayor error en la estimación	mayor error en la estimación
Diferenciación Automática (AD Model Builder)	presenta los mejores resultados	presenta los mejores resultados

Cuadro 5.3: Cuadro comparativo de los resultados obtenidos por cada método utilizado en el modelo Stock-Reclutamiento.

#### 5.1.5. Modelo estadístico de evaluación del stock basado en captura por edad y por tallas

En esta sección se realizó una aplicación del método de optimización *diferenciación automática* y el método de estimación *máxima verosimilitud*.

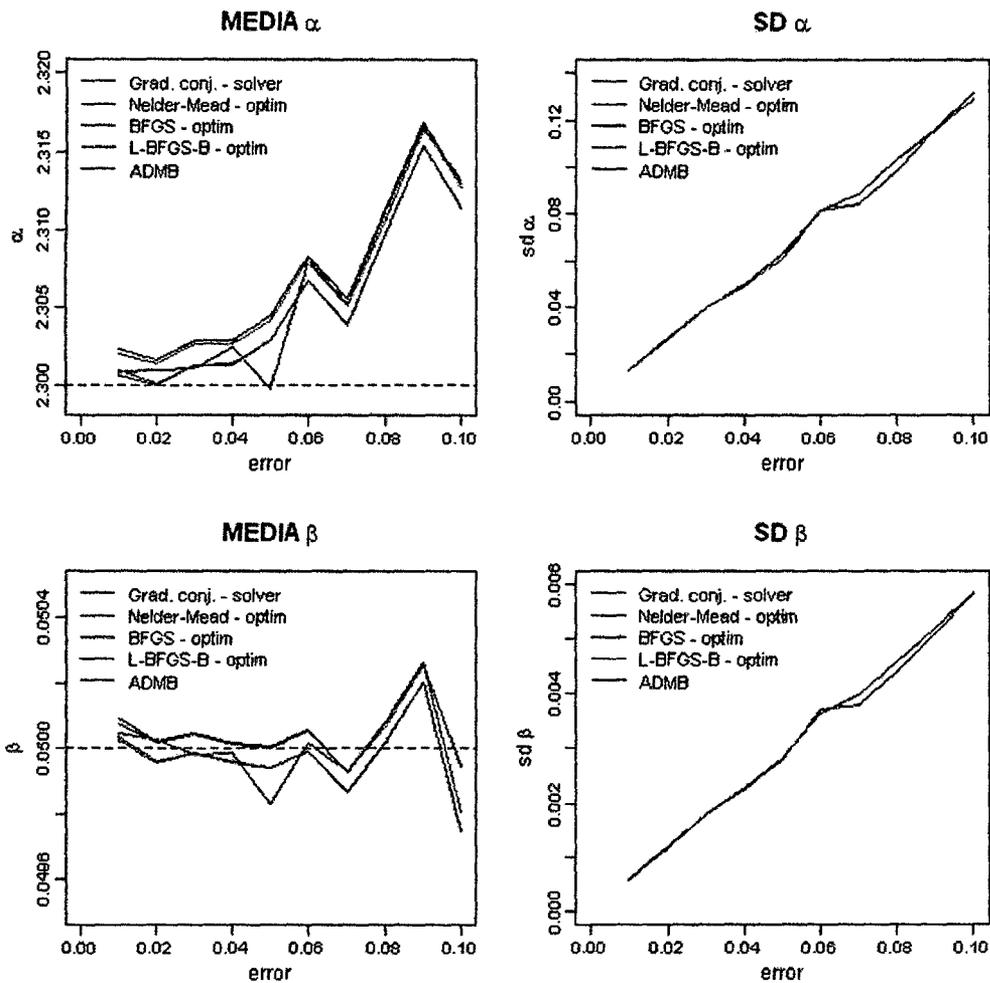


Figura 5.6: Resultados de las estimaciones de los parámetros del modelo Stock-Reclutamiento, para datos con diferente grado de dispersión. La línea discontinua roja indica el valor real del parámetro.

En este caso, se utilizó el método de evaluación adoptado por la *Organización Regional de Ordenamiento Pesquero del Pacífico Sur* para el recurso jurel, llamado *Joint Jack Mackerel (JJM)*.

Los datos ingresados fueron los desembarques anuales de los años 1970-2012 (Figura 5.7) y las estimaciones de biomasa acústica 1970-2012 (Figura 5.8) y captura por unidad de esfuerzo (CPUE) 2002-2012 (Figura 5.9).

Se realizaron tres simulaciones manteniendo siempre los datos de desembarques, y utilizando los datos de biomasa acústica y CPUE alternadamente.

Consideramos Modelo 1 al utilizar los datos de biomasa acústica y CPUE, Modelo 2 a la simulación considerando solo los datos de biomasa acústica y Modelo 3 a la simulación considerando solo la CPUE.

Los resultados obtenidos muestran variaciones según los datos utilizados en cada modelo. Las estimaciones de biomasa desovante (Figura 5.10), reclutamiento (Figura 5.11) y biomasa total (Figura 5.12) fueron calculadas con intervalos de confianza en función de la desviación estándar de los valores resultantes. También se calculó la mortalidad por pesca (Figura 5.13) y el número de individuos (Figura 5.14).

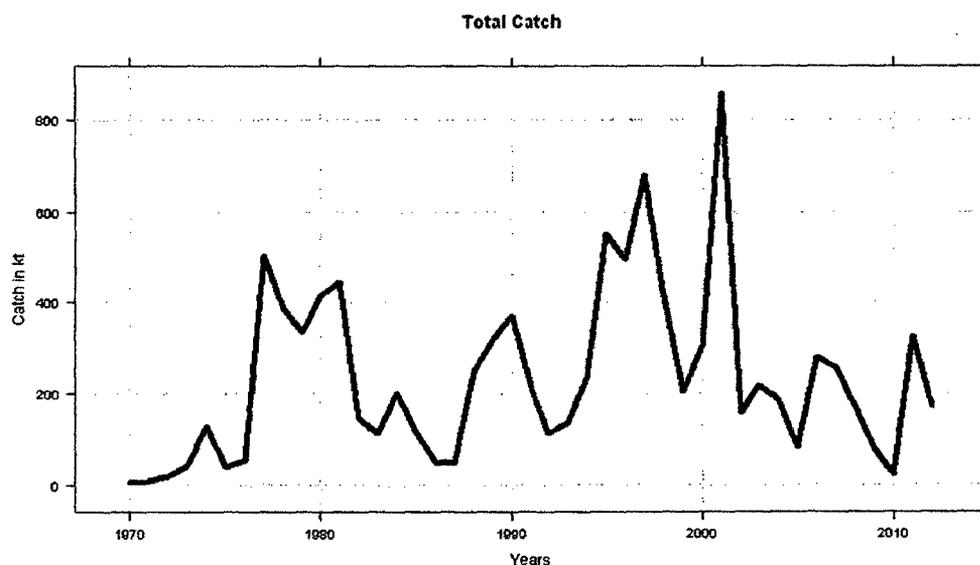


Figura 5.7: Desembarque anual de jurel durante el periodo 1970-2012

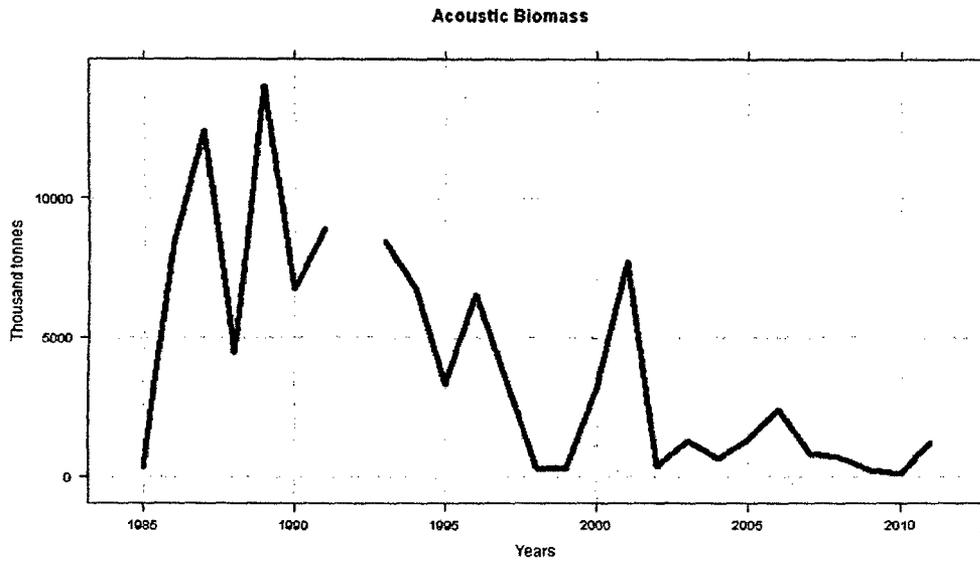


Figura 5.8: Biomasa acústica estimada durante el periodo 1970-2012

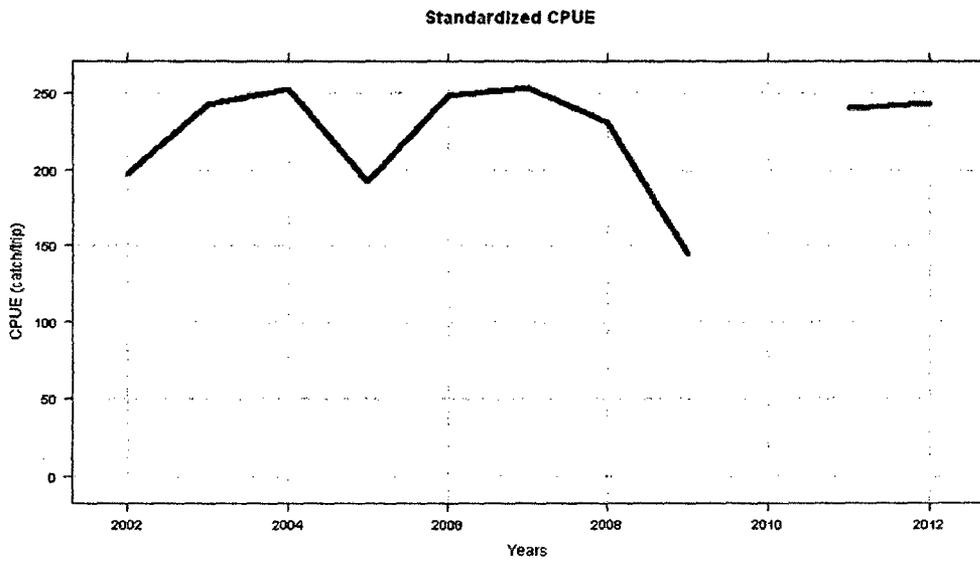


Figura 5.9: Captura por unidad de esfuerzo (CPUE) estandarizada durante el periodo 2002-2012

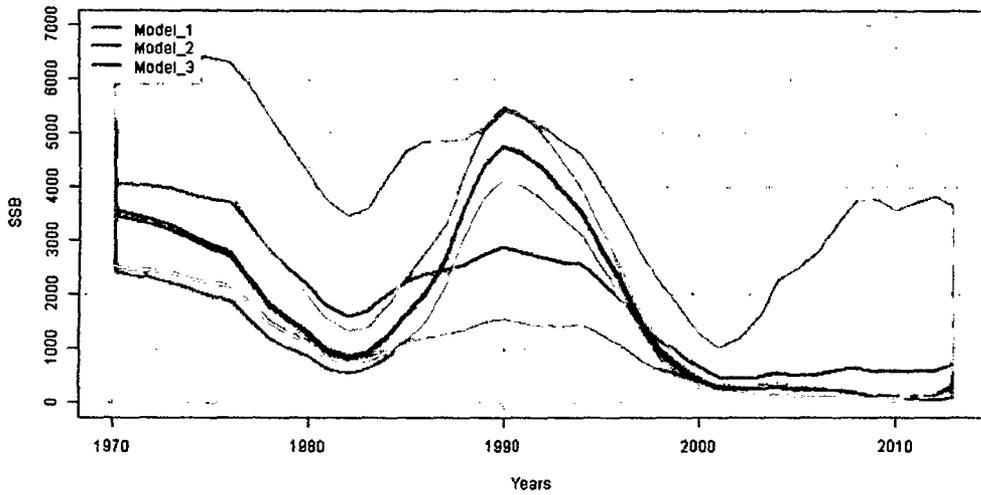


Figura 5.10: Estimación del stock de biomasa desovante del jurel (1970-2012) calculado con los tres modelos estudiados

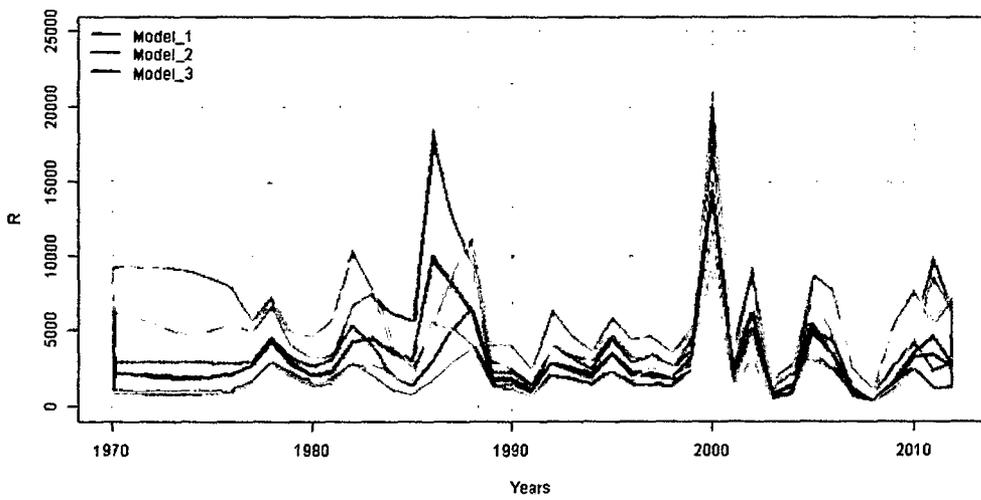


Figura 5.11: Estimación del número de reclutas para el jurel (1970-2012)

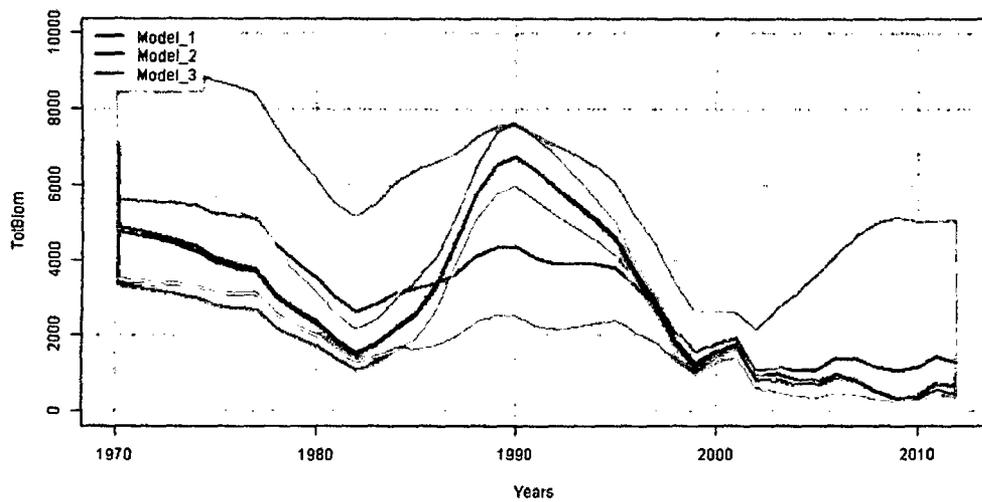


Figura 5.12: Estimación de la biomasa total de jurel (1970-2012)

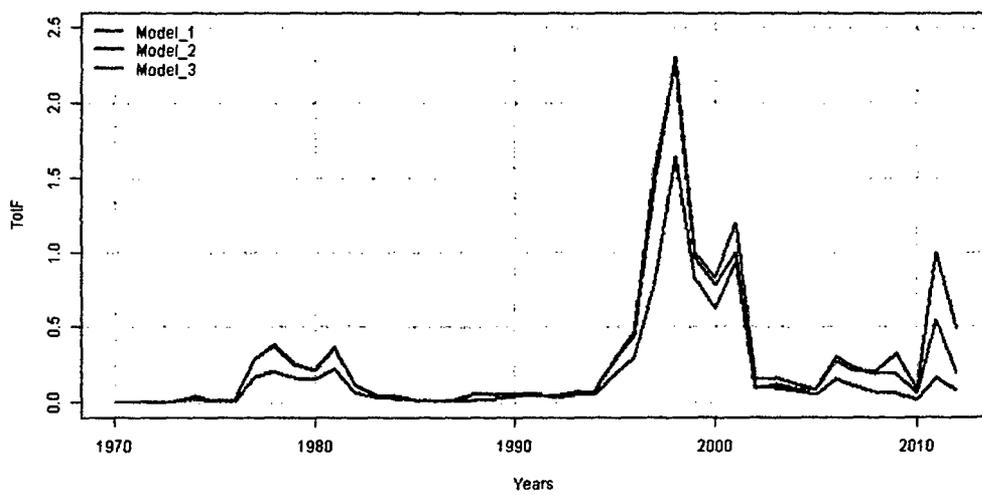


Figura 5.13: Estimación de la mortalidad por pesca del jurel (1970-2012)

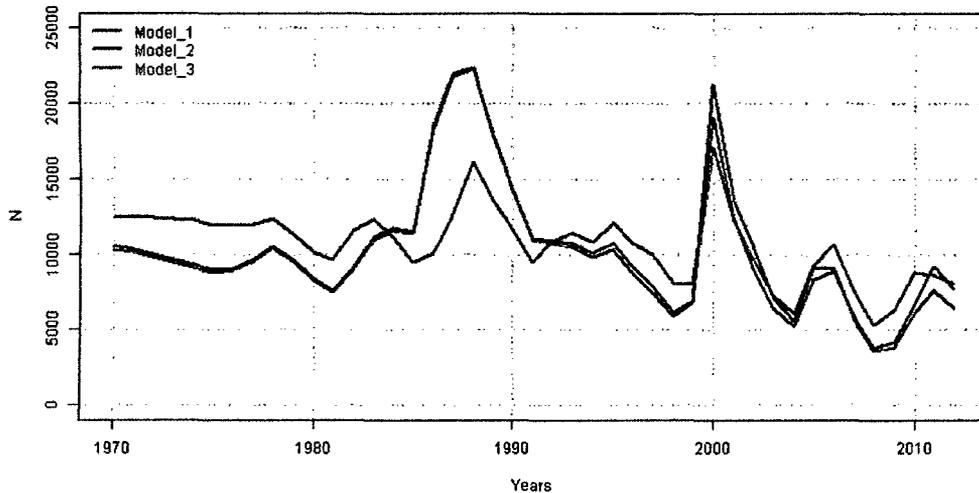


Figura 5.14: Estimación del número de individuos de jurel (1970-2012)

## 5.2. Discusión

En este estudio se compararon los resultados obtenidos con cada modelo utilizado, encontrando algunas diferencias en los resultados.

Las principales diferencias encontradas se deben al requisito de diferenciabilidad que exigen los métodos.

Otro factor que determinó que el modelo devuelva un valor alejado del real se debe a la variabilidad de los datos producto de agregarle un error log-normal.

En algunos casos, dependiendo de los parámetros de entrada utilizados o de los datos utilizados, el modelo no era capaz de ejecutarse.

Para cada caso o modelo estudiado obtenemos un método de optimización adecuado:

- Los mejores métodos fueron *Nelder-Mead* y *diferenciación automática*. El primero es debido a la regularidad de la función del modelo, el segundo es porque dicho método utiliza derivadas exactas en vez de un cálculo numérico, minimizando así el error de proceso.
- El método de diferenciación automática implementado en *ADMB* fue el que dio consistentemente los mejores resultados.

- Entre los métodos de estimación utilizados se prefiere el método de *máxima verosimilitud*, pues este hace el proceso de minimización más simple.
- El mejor proceso de regularización fue restringir el espacio del dominio al acotar los parámetros de entrada.

En este estudio hemos visto la importancia que tiene el método de estimación y el método de optimización al momento de resolver un problema.

El error de proceso, producido por el método de optimización, depende no solamente de la elección del método, sino también de la herramienta con la que se va a trabajar. Este error presenta una limitación, por lo general no se puede manipular a fin de minimizarlo.

Un gran número de softwares han sido aplicados al análisis cuantitativo de la dinámica de poblaciones de peces. Estos incluyen lenguajes de programación tradicionales, programas comunes de hojas de cálculo y lenguajes de cuarta generación que se ocupan de problemas específicos para el análisis cuantitativo de la dinámica de poblaciones de peces (Maunder et al., 2009).

Programas como hojas de cálculo son más atractivas y mayormente utilizadas por personas que tienen poca experiencia en programación debido al diseño intuitivo de los cálculos y la visualización de los resultados, sin embargo, pueden ser propensos a errores de programación, y su código es difícil de documentar, revisar y mantener (Prager & Williams, 2003).

El caso de la herramienta *solver* de *Excel* permite la estimación de parámetros, sin embargo, el algoritmo de utiliza algunas veces no funciona adecuadamente en problemas mal condicionados con muchos parámetros.

En general, el usar hojas de cálculo en aspectos de análisis de pesquerías es efectivo para fines inmediatos, sin embargo, su uso en la evaluación de poblaciones se vuelve más complejo y se presentan dificultades en la elaboración. La estimación de la incertidumbre y los tiempos de ejecución para modelos complejos, tales como la estimación de parámetros no lineales) son algunas de las dificultades que se presentan (Maunder et al., 2009).

Actualmente, AD Model Builder juega un papel importante como un entorno de software para estimar los parámetros de modelos de evaluación de poblaciones pesqueras complejos y altamente parametrizados.

Maunder et al. (2009) nos dice que el método “reverso” que utiliza el *ADMB* actúa de manera más rápida y exacta que un cálculo de derivadas

numéricas basada en diferencias finitas, y con este estudio podemos afirmar que el *ADMB* es el mejor método a utilizar en la resolución de modelos de recursos pesqueros, frente a otros que utilizan los métodos de *Newton* o *Quasi-Newton*.

El error de observación, producido por el método de estimación, es más flexible, dependiendo del problema que se quiere resolver se puede “elegir” el error, ya sea por la distribución de los datos de entrada como las facilidades de las herramientas y costos de estimación.

La elección del método de estimación y optimización depende del problema a resolver, al igual que los datos con los que se dispone.

De las simulaciones realizadas para los modelos longitud-peso, talla-edad y stock-reclutamiento, notamos que las realizadas con el software *ADMB* se aproximaban más a los valores reales, lo cual indica que el método de optimización “*diferenciación automática*” es el método que mejor ajusta en todos los casos.

Sabiendo lo anteriormente expresado, procedimos a aplicar los métodos seleccionados a un problema real.

La evaluación del stock del jurel es un problema que en los últimos años ha tomado mayor importancia en todo el Pacífico Sur debido a que se ha evidenciado una disminución en los desembarques. En la reunión de la OROP se ha formado un subgrupo científico para velar por la conservación de este recurso.

Una medida adoptada comprende hacer la evaluación del stock y proyecciones para todo el Pacífico Sur, siguiendo el modelo *Joint Jack Mackerel* (JJM), el cual está implementado en el software *AD Model Builder*.

Siguiendo esa línea, se han realizado evaluaciones para la región norte (la cual comprende Perú).

Este ejercicio comprende lo visto en la sección del modelo estadístico de captura por talla, observándose la importancia no sólo de los métodos de optimización y de estimación elegidos, sino también la importancia de los datos de entrada.

Al cambiar los datos de entrada (cantidad de datos o tipo de datos), varía la función objetivo, y por lo tanto, también varían las estimaciones y la optimización, convirtiéndose así en un nuevo problema y un nuevo modelo.

Para apreciar las diferencias introducidas al considerar distintos tipos de

datos, se procedió a correr el modelo JJM para la región peruana con series de datos de biomasa acústica, series de captura por unidad de esfuerzo y una combinación de estos.

### **5.3. Perspectivas**

Aplicar los resultados a otros modelos en dinámica de poblaciones para la evaluación de recursos pesqueros dentro del territorio peruano, implementar regularización bayesiana.

# Referencias

- Aster, R. C., Borchers, B., & Thurber, C. H. (2005). *Parameter Estimation and Inverse Problems*, volume 90 of *International Geophysics Series*. Elsevier.
- Bard, Y. (1974). *Nonlinear Parameter Estimation*. New York: Academic Press.
- Bertsekas, D. P. (1999). *Nonlinear Programming* (Second ed.). Optimization and Computation Series. Massachusetts: Athena Scientific.
- Bolker, B. M., Gardner, B., Maunder, M. N., Berg, C. W., Brooks, M., Comita, L., Crone, E., Cubaynes, S., Davies, T., de Valpine, P., Ford, J., Gimenez, O., Kéry, M., Kim, E. J., Lennert-Cody, C., Magnusson, A., Martell, S., Nash, J. C., Nielsen, A., Regetz, J., Skaug, H. J., & Zipkin, E. (2013). Strategies for fitting nonlinear ecological models in R, AD Model Builder, and BUGS. *Methods in Ecology and Evolution*, 4, 501–512.
- Broyden, C. G. (1970a). The convergence of a class of double-rank minimization algorithms. i: General considerations. *Journal of the Institute of Mathematics and its Applications*, 6, 76–90.
- Broyden, C. G. (1970b). The convergence of a class of double-rank minimization algorithms. ii: The new algorithm. *Journal of the Institute of Mathematics and its Applications*, 6, 222–231.
- Byrd, R. H., Lu, P., Nocedal, J., & Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16(5), 1190–1208.

- Conn, A. R., Scheinberg, K., & Vicente, L. N. (2009). *Introduction to Derivative-Free Optimization*. MPS-SIAM Series on Optimization. Philadelphia: SIAM.
- Crouzeix, J. P., Keraghel, A., & Sosa, W. (2011). *Programación Matemática Diferenciable*. Lima. Perú: Universidad Nacional de Ingeniería.
- Deriso, R. B. (1980). Harvesting strategies and parameter estimation for an age-structured model. *Canadian Journal of Fisheries and Aquatic Sciences*, 37, 268–282.
- Deriso, R. B., Maunder, M. M., & Skalski, J. R. (2007). Variance estimation in integrated assessment models and its importance for hypothesis testing. *Canadian Journal of Fisheries and Aquatic Sciences*, 64(2), 187–197.
- Dowd, M. & Meyer, R. (2003). A bayesian approach to the ecosystem inverse problem. *Ecological Modelling*, 168, 39–55.
- Fletcher, R. (1970). A new approach to variable metric algorithms. *The Computer Journal*, 13(3), 317–322.
- Fletcher, R. (2000). *Practical Methods of Optimization* (Second ed.). A Wiley-Interscience Publication. New York: Wiley.
- Fournier, D. A., Skaug, H. J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M. N., Nielsen, A., & Sibert, J. (2012). AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optimization Methods and Software*, 27(2), 233–249.
- Gallucci, V. F. & Quinn, II, T. J. (1979). Reparameterizing, fitting, and testing a simple growth model. *Transactions of the American Fisheries Society*, 108(1), 14–25.
- Gasulla Forner, M. (1999). *Obtención de imágenes de la distribución de impedancia eléctrica en el subsuelo. Aplicación a la detección de objetos locales*. Tesis doctoral, Universitat Politècnica de Catalunya. Departament d'Enginyeria Electrònica.

- Goldfarb, D. (1970). A family of variable-metric methods derived by variational means. *Mathematics of Computation*, 24(109), 23–26.
- Griewank, A. & Walther, A. (2008). *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation* (Second ed.). Philadelphia: SIAM.
- Hadamard, J. (1902). Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin*, 13, 49–52.
- Hilborn, R. & Walters, C. J. (1992). *Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty*. Dordrecht: Springer.
- ICES (2012). Report on the classification of stock assessment methods developed by SISAM. Technical report, ICES CM 2012/ACOM/SCICOM:01.
- Izmailov, A. & Solodov, M. (2007). *Otimização, Volume 2: Métodos computacionais*. Rio de Janeiro: IMPA.
- Lange, K. (2004). *Optimization*. New York: Springer-Verlag.
- Maunder, M. N. & Punt, A. E. (2013). A review of integrated analysis in fisheries stock assessment. *Fisheries Research*, 142, 61–74.
- Maunder, M. N., Schnute, J. T., & Ianelli, J. N. (2009). Computers in fisheries population dynamics. In B. A. Megrey & E. Moksness (Eds.), *Computers in Fisheries Research* (Second ed.). Dordrecht: Springer.
- Menke, W. (1989). *Geophysical Data Analysis: Discrete Inverse Theory*. International Geophysics Series. New York: Academic Press.
- Millar, R. B. (2011). *Maximum likelihood estimation and inference. With examples in R, SAS and ADMB*. Chichester, UK: John Wiley & Sons, Ltd.
- Monett, D. & Burkhard, H.-D. (2004). Problemas inversos, técnicas evolutivas y agentes inteligentes: Estrechando las fronteras. *Avances en Sistemas e Informatica*, 1, 29–34.
- Mosegaard, K. & Tarantola, A. (1995). Monte Carlo sampling of solutions to inverse problems. *Journal of Geophysical Research*, 100(B7), 12431–12447.

- Nash, J. C. (1990). *Compact Numerical Methods for computers. Linear Algebra and Function Minimisation* (Second ed.). New York: Adam Hilger.
- Nelder, J. A. & Mead, R. (1965). A simplex method for function minimization. *The Computer Journal*, 7(4), 308–313.
- Nocedal, J. & Wright, S. J. (2006). *Numerical Optimization* (Second ed.). Springer Series in Operations Research and Financial Engineering. New York: Springer.
- Pauly, D., Muck, P., Mendo, J., & Tsukayama, I. (1989). *The Peruvian Upwelling Ecosystem: Dynamics and Interactions*. ICLARM Conference Proceedings 18. IMARPE.
- Pauly, D. & Tsukayama, I. (1983). On the seasonal growth, monthly recruitment and monthly biomass of peruvian anchoveta (*engraulis ringens*) from 1961 to 1979. In G. D. Sharp & J. Csirke (Eds.), *Proceedings of the Expert Consultation to examine changes in abundance and species of neritic fish resources*, FAO Fisheries Report No. 291, Vol. 3 (pp. 987–1004). San José. Costa Rica.
- Pauly, D. & Tsukayama, I. (1987). *The Peruvian Anchoveta and Its Upwelling Ecosystem: Three Decades of Change*. ICLARM Studies and Reviews 15. IMARPE.
- Prager, M. H. & Williams, E. H. (2003). From the golden age to the new industrial age: fishery modeling in the early 21st century. *Natural Resource Modeling*, 16(4), 477–489.
- Quinn, II, T. J. & Deriso, R. B. (1999). *Quantitative Fish Dynamics*. New York: Oxford University Press.
- R Development Core Team (2012). *R: A Language and Environment for Statistical Computing*. Vienna, Austria. ISBN 3-900051-07-0.
- Ramos-Vásquez, E. (2013). Protocolo: Estimación de la captura total permisible del stock peruano de jurel. Technical report, Instituto del Mar del Perú. Lima, Perú.

- Ricker, W. E. (1975). Computation and interpretation of biological statistics of fish populations. *Bulletin of the Fisheries Research Board of Canada*, 191.
- Shanno, D. F. (1970). Conditioning of quasi-newton methods for function minimization. *Mathematics of Computation*, 24(111), 647–656.
- Sóbol, I. (1976). *Método de Montecarlo*. Moscú: Editorial MIR.
- SPRFMO (2012). Report of the 11th Science Working Group - Annex SWG-03. Technical report, Lima, Perú.
- SPRFMO (2013). Report of the 1st Scientific Committee Meeting - Annex 05. Technical report, La Jolla, USA.
- Tarantola, A. (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation*. Philadelphia: SIAM.
- Tarantola, A. (2006). Popper, bayes and the inverse problem. *Nature Physics*, 2, 492–494.
- Walter, E. & Pronzato, L. (1997). *Identification of Parametric Models from Experimental Data*. Great Britain: Springer.