Revista EIA, ISSN 1794-1237 / Year XIII / Volume 13 / Issue N.25 / January-June / pp. 81-98
Technical-scientifical biannual publication / Universidad EIA, Envigado (Colombia)

ESTIMATION OF PARAMETERS IN NONLINEAR MODELS: ALGORITHMS AND APPLICATIONS

Óscar Cornejo Zúñiga[1]
Rodrigo Rebolledo Vega[1]

ABSTRACT

This article shows different algorithms for estimating parameters in nonlinear models. They are applied primarily to a database of problems classified as difficult. Later, the article shows the behavior of the algorithms for the study of growth in the anchovy and sardine, and males and females of the common hake by adjusting a Von Bertalanffy model. The Cerrato test is applied for growth comparisons between sexes for the common hake. The algorithms are implemented in a MATLAB environment, showing good behavior regarding CPU time, number of iterations and exactitude in the solution found with respect to certified values of the problems in the database.

KEYWORDS: Nonlinear regression; Quasi-Newton methods; Confidence region methods; Growth models.

ESTIMACIÓN DE PARÁMETROS EN MODELOS NO LINEALES: ALGORITMOS Y APLICACIONES

RESUMEN

En este artículo se muestran diferentes algoritmos para estimar parámetros en modelos no lineales. Se aplican primeramente a una base de datos de problemas clasificados difíciles. Posteriormente, se muestra el comportamiento de los algoritmos para el estudio de crecimiento de la merluza común en machos y hembras, anchoveta y sardina común ajustando un modelo de Von Bertalanffy. Se aplica el test de Cerrato para la comparación de crecimientos entre géneros para la merluza común. Los algoritmos se implementaron en ambiente MATLAB presentando un buen comportamiento en cuanto a tiempo CPU, número de iteraciones y exactitud de la solución encontrada respecto de valores certificados de los problemas de la base de datos.

PALABRAS CLAVE: Regresión no-Lineal; Métodos Cuasi-Newton; Métodos de Región de Confianza; Modelos de Crecimiento.

[1] Industrial Engineering Department, College of Engineering. Universidad Católica de la Santísima Concepción, Concepción, Chile.

Author of correspondence: Rebolledo-Vega, R. (Rodrigo). Alonso de Ribera 2850, Concepción, Chile / Tel.: 56-41-2345548
Email: rrebolledo@ucsc.cl

Paper history:
Paper received: 02-III-2015/ Approved: 04-III-2016
Available online: October 30 2016
Open discussion until October 2017

DOI: http:/dx.doi.org/10.14508/reia.2016.13.25.81-98

ESTIMATIVA DE PARÂMETROS EM MODELOS NÃO LINEARES: ALGORITMOS E APLICAÇÕES

RESUMO

Neste artículo se mostram diferentes algoritmos para estimar parâmetros em modelos não lineares. Aplica-se primeiro a um banco de dados de problemas classificadas difíceis. Subsequentemente, se mostra o comportamento de algoritmos para o estudo do crescimento da Pescada comum em machos e fêmeas, anchova e sardinha comum ajustando um modelo de Von Bertalanffy. Que aplica o teste de Cerrato para a comparação do crescimento entre gêneros para a pescada comum. Os algoritmos se programaram em ambientes MATLAB apresentando um bom desempenho em termos de tempo de CPU, número de iterações e precisão da solução encontrados valores certificados dos problemas do banco de dados.

PALAVRAS-CHAVE: Regressão não linear; Métodos Quasi-Newtown; Métodos da região de confiança; modelos de crescimento.

1. INTRODUCTION

In the majority of the studies that involve data it is necessary to conduct estimations. For example, to estimate demand for certain products, to estimate parameters in biological growth models, in chemical reactions, etc. The applications are extremely varied in all spheres of knowledge: engineering sciences, social sciences, medicine, ecology, botany, political science, finance, among others. These estimations are studied through models that depend on a joining of experimental parameters and data. We must determine the best model under a certain criterion, generally least squares. Therefore, the desire exists to know about algorithms that possess certain convergence properties that permit, in a reasonable calculation time, parameter estimation, especially if the model studied is nonlinear in its parameters. In this article, we will show some algorithms and their principal characteristics that allow us to estimate parameters in nonlinear models. The article is organized in the following manner: in section 2 we study the most relevant aspects of algorithms utilized in the general problem of parameter estimation. Section 3 shows the performance of the algorithms on a database of problems categorized as difficult. Section 4 shows a study on the fishing sector in which the aim is to estimate the growth parameters for anchovies, the common sardine and the common hake. The importance of this application lies in being able to scientifically predict certain aspects related to the exploitation of this resource and the possibility of understanding spawning, catch or recruitment seasons, which are of vital importance for sensible management of marine resources. Finally, in section 5 some conclusions about this project are given.

2. GENERAL SHAPE OF NONLINEAR REGRESSION MODELS

Nonlinear regression models are very similar in their general form to linear regression models. Each observation y_i is written in terms of the nonlinear response function $f(x_i; \theta)$ and a random error term of ε_i. For the error term ε_i it is assumed that it is a normal independent variable σ^2 with the variance (Draper & Smith, 1966; Neter *et al.*, 1996). An important difference in nonlinear regression models is that the number of regression parameters θ is not directly related to the number of variables x_i in the model. The general form of a nonlinear regression model will be the following:

$$y_i = f\,(x_i;\theta)+\varepsilon_i;\varepsilon\sim N(0,\sigma^2) \quad (1)$$

where $x_i(m\times1) = [x_{i\,1}\ x_{i\,2} \cdots x_{i\,m}]$; $\theta(n\times1) = [\theta1\ \theta2 \cdots \theta n]$ correspond to the data vector and parameters respectively, and $\sim N\ (0,\ \sigma^2)$ is the distribution of normal probabilities with median 0 and variance σ^2. To determine the parameters, we solve the following optimization problem: *minimize the sum of squared errors*, e_i, where we define $e_i = y_i - f(x_i, \theta)$ as the i^{th} error term

$$S(\theta)= \sum_{i=1}^{m} [y_i-f(x_i,\theta)]^2 = \left\|y-f(\theta)\right\|_2^2 \qquad (2)$$

with $y = (y_1, y_2,..., y_m)^t$; $f_i(\theta) = f(x_i,\theta)$; $f(\theta) = (f_1(\theta), f_2(\theta)..., f_m(\theta))^t$. Note that (2) corresponds to an unrestricted optimization problem, which can be solved by any general optimization method; see, for example, De la Fuente O'Conor (1995), Fletcher (1980), Nocedal & Wright (1999). However, given the particular shape of (2), methods have been created for exploiting the special structure of this type of problem. The Jacobian matrix J of $S(\theta)$ is given by

$$J(\theta)= \frac{\partial f(\theta)}{\partial \theta^t} =[\nabla f_1(\theta)\ \nabla f_2(\theta)...\nabla f_m(\theta)]^t \in \mathbb{R}^{m\times n} \qquad (3)$$

Linearizing $f_i(\theta)$ at point one θ_k has the linear system

$$f(\theta_k)+J(\theta_k)\ d=0 \qquad (4)$$

With $d= \theta-\theta_k$. In the case of $m= n$ y $J(\theta_k)$ non-singular and the previous linear system brings us to the **Newton-Raphson** method

$$\theta_{k+1}= \theta_k-J(\theta_k)^{-1}f(\theta_k) \qquad (5)$$

If $m \neq n$ or $J(\theta_k)$ is singular, (4) might not have solutions. Therefore, it is only natural to replace (4) with the problem of linear least squares

$$\min_{d\in \mathbb{R}^n} \left\|f(\theta_k)+J(\theta_k)d\right\|_2^2 \qquad (6)$$

which can be considered a linearization of (2). The minimum norm solution in the previous sub-problem is the Gauss-Newton step

$$d_k^{GN}=-(J(\theta_k))^{\dagger} f(\theta_k) \qquad (7)$$

where $(J(\theta_k))^{\dagger}$ is the Moore-Penrose generalized inverse of $J(\theta_k)$. If $J(\theta_k)$ is the complete column range, then the Gauss-Newton direction is written as

$$d_k^{GN}=-((J(\theta_k)^t J(\theta_k))^{-1} J(\theta_k)^t f(\theta_k) \qquad (8)$$

by obtaining the **Gauss-Newton** (G-N) method.

$$\theta_{k+1}= \theta_k-((J(\theta_k)^t J(\theta_k))^{-1} J(\theta_k)^t f(\theta_k),\ \forall k \qquad (9)$$

Note that in this case we have considered a step size $\alpha_k=1, \forall k$. However, in the previous iteration we can consider a variable step size in each iteration. For that reason, a step size of α_k is considered given by Armijo, Goldstein, Wolfe or Thuente. See, for example, Nocedal & Wright (1999) which provides the origin of the **dampedGauss-Newton** method.

2.1. Quasi-Newton methods

Quasi-Newton methods consist of approximating the Hessian matrix of each iteration by way of recurrence formulas that relate to the value it takes in preceding iterations, see Bonnans *et al.* (2002). The search direction in the Newton method requires the calculation of the Hessian matrix and that it be inversible, which cannot be guaranteed in the course of the iterations. This entails a great force from the computational point of view in the calculation of this matrix. With the aim of getting around these difficulties the Quasi-Newton methods approximate the matrix $\nabla^2 f(\theta_k)$ by a matrix defined B positive, which is modified in each iteration and converges at the true Hessian matrix, see Coleman (1984), Frandsen *et al*. (2004), Lange (2004). The Quasi-Newton methods have proved to be quite efficient in nonlinear optimization and play an important role in many implementations. Furthermore, this type of method, as opposed to those of Newton, has a superlinear convergence rate, which frequently, from a computational standpoint, is more efficient in the end than the Newton analytical method, see De la Fuente O'Conor (1995), Luenberger (1984). In these methods, the iterations can be more computationally costly; however, the information stored in the Hessian approximation may be able to reduce the total number of iterations compared with other traditional methods (Nocedal & Wright, 1999). Let's consider the solution of the system

$$B_k d_k=-\nabla f(\theta_k) \qquad (10)$$

where B_k is a positive defined square matrix. Another way to present the quasi-Newton methods is through the approximation of the inverse of the Hessian, meaning $B = H^{-1}$. Like any iterative method and according to what was previously mentioned, this depends upon an initial approximation. In this case an approximation is also needed for the Hessian, or B_0 initial, which frequently can be taken as the identity matrix $B_0=I$ if more information does not exist. In addition, B_0 can be considered a multiple of the identity matrix, meaning $B_0 = \eta I$, for a $\eta > 0$. The Hessian matrix is updated based on the following structure (Nocedal y Wright, 1999):

$$B_{k+1} = B_k + U_k, \quad k = 0,1,... \tag{11}$$

where U_k is the expression that approximates the true Hessian matrix. Let's look at two possible strategies. One condition to define B_k is, looking Frandsen *et al*. (2004), Luenberger (1984), Nocedal & Wright (1999):

$$B_{k+1}(\theta_{k+1}-\theta_k)=\nabla f(\theta_{k+1}-\nabla f(\theta_k)) \tag{12}$$

This condition is known as the *secant condition*, which is based on a generalization of the one-dimensional secant method, where the Hessian matrix $\nabla^2 f(\theta_k)$ is replaced by an approximation B_k. Defining $S_k = \theta_{k+1} - \theta_k = \alpha_k d_k$ and $y_k = \nabla f(\theta_{k+1}) - \nabla f(\theta_k)$ one obtains $B_{k+1}S_k = y_k$. The secant condition (Kelley, 1995) is satisfied if:

$$S_k^T \; y_k > 0 \tag{13}$$

which is known as the curvature condition[1] (Frandsen *et al*., 2004). The update matrix B_k can be calculated by different methods. Below are two methods for updating said expression. This method was developed by Broyden, Fletcher, Goldfard and Shanno, known as **BFGS**, and takes the following form. See Fletcher (1980), Frandsen *et al.* (2004), Luenberger (1984):

$$U_k = -\frac{(B_k S_k)(B_k S_k)^T}{S_k^T B_k S_k} + \frac{y_k y_k^T}{y_k^T S_k} \quad k = 0, 1, ... \tag{14}$$

One of the most intelligent schemas for the construction of the inverse of the Hessian was originally proposed by Davidon and later developed by Fletcher and Powell, known today as **DFP**. The update is given by (ver Fletcher (1980), Frandsen *et al.* (2004), Luenberger (1984));

$$U_k = H_k + \frac{S_k S_k^T}{S_k^T y_k} - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k}, \quad k = 0, 1, ... \tag{15}$$

2.2. Trust region

Trust region methods offer a framework for guaranteeing algorithm convergence. They were first used to solve nonlinear least squares problems and later were adapted for more general optimization problems. These methods explicitly make reference to a **model** of the objective function. In the Newton method, this model is quadratic and is obtained from the Taylor series of f around point x_k. The method will "believe" in this model only in the vicinity of point θ_k, defined by the restriction $\|p\| \leq \Delta_k$. This will serve to limit the size of the step from θ_k to θ_{k+1}. The value of Δ_k is adjusted based on agreement in the model

$$\psi_k(p)=f(\theta_k)+\nabla^t f(\theta_k)p+\frac{1}{2}p^t \nabla^2 f(\theta_k)p \tag{16}$$

and the objective function $f(\theta_k+p)$. If the agreement is good, the model is reliable and is increased Δ_k. If not, Δ_k is decreased. In iteration k of a trust region method, the following subproblem is solved:

$$\min_p \; \psi_k(p)=f(\theta_k)+\nabla^t f(\theta_k)p+\frac{1}{2} p^t \nabla^2 f(\theta_k)p \;\; s/a: \|p\| \leq \Delta_k \tag{17}$$

which corresponds to a restricted problem. The optimality conditions show that p_k will be the solution of the linear system:

$$(\nabla^2 f(\theta_k) + \lambda I)\, p_k = -\nabla f(\theta_k)$$

with $\lambda>0$, $(\nabla^2 f(\theta_k)+\lambda I)$ is defined as positive, and

$$\lambda(\Delta_k - \|p_k\|) = 0.$$

If $\nabla^2 f(\theta_k)$ is defined as positive and Δ_k is sufficiently large, the solution to the subproblem is

1 The condition of the secant always has a solution if the curvature condition is valid.

$$\nabla^2 f(\theta_k)p = -\nabla f(\theta_k)$$

the solution from the Newton equations. The method depends on the reason; see Borlin (2007), Madsen *et al.* (2004), Mizutani (1999);

$$\rho_k = \frac{f(\theta_k)-f(\theta_k+p_k)}{\psi_k(0)-\psi_k(p_k)} = \frac{\text{current reduction}}{\text{planned reduction}} \quad (18)$$

If ρ_k is negative the step must be refused; furthermore, if ρ_k is close to 1, agreement exists between the model and the objective function, so it is safe to extend the trust region for another iteration. If ρ_k is positive but not close to 1, the region is not altered. Conversely, if ρ_k is close to 0 or negative the trust region is decreased. It is important to point out that the solution to the subproblem does not have to be exact; one can approximate by way of the Cauchy point, or the Dogleg or Steihaug methods, among others. See Borlin (2007), Coleman (1984), Madsen *et al*. (2004), Mizutani (1999).

2.3. Levenberg-Marquardt

The Levenberg-Marquardt method (L-M) has been a standard technique for nonlinear least squares problems, commonly used in various disciplines for adjusting data. This iterative algorithm can be seen as a combination of the methods of maximum descent and the Gauss-Newton method. When the current solution is found to be far from the local minimum, the algorithm behaves like the maximum descent method: slow but with guaranteed convergence. However, when the solution is close to the local minimum the method exhibits behavior similar to Gauss-Newton with a rapid convergence. L-M has emerged as a good alternative for avoiding the problems that the Gauss-Newton method presents when the Jacobian matrix is not singular. Levenberg (1944) and Marquardt (1963) suggested calculating the direction $p_k=\theta-\theta_k$, using the solution to the following problem. See Borlin (2007), Coleman (1984), De la Fuente O'Conor (1995), Madsen *et al.* (2004), Nocedal & Wright (1999);

$$\min_{p_k \in R^n} \{\|f(\theta_k)+J(\theta_k)p_k\|_2^2 = \mu_k\|p_k\|_2^2\} \quad (19)$$

We must note that the parameter μ_k controls the size of the vector p_k. Observe further that p_k is defined as inclusive if $J(\theta_k)$ is not a complete range. Under, $\mu_k \to \infty$, $\|p_k\| \to 0$ and p_k it becomes parallel to the maximum slope. The L-M method can be described and analyzed under the framework of the trust region methods (in effect, this method is considered a precursor of the trust region methods for unrestricted optimization). See Mizutani (1999). hus defined, this subproblem can be proven equivalent to the following optimization problem:

$$\min_{p \in R^n} \|f(\theta_k)+J(\theta_k)p\|_2 \ \ s.a. \ \|p\|_2 \le \Delta_k \quad (20)$$

where $\Delta_k > 0$ is the radius of the trust region. In effect, the model function $m_k(p)$ will be:

$$m_k(p) = \frac{1}{2}\|f_k\|^2 + p^T J_k^T f_k + \frac{1}{2} p^T J_k^T J_k p$$

It is known that when the Gauss-Newton solution p_{GN} falls within the limits of the trust region, which is to say $\|p_{GN}\| < \Delta$, it can be considered a solution to the subproblem. Furthermore, a scale exists $\lambda > 0$ just like the solution $p=p_{LM}$ satisfies $\|p\| = \Delta$ and

$$(J^T J + \lambda I)\, p = -J^T f \quad (21)$$

which corresponds to the update of the step to be considered; See Berghen (2004).

3. EVALUATION PROBLEMS

In this section, we validate the G-N, damped G-N, BFGS, DFP, and L-Malgorithms using problems of a high grade of difficulty known as Thurber, BoxBOD, Rat42, Bennett5, Rat43 and Eckerle4, taken from the Statistical Reference Datasets Project (STRD), developed by the staff at the Statistical Engineering Division and the Mathematical and Computational Sciences Division of the National Institute of Standards and Technology (NIST), which provides reference databases with certified values.

For each of these problems the function to minimize, the parameters to estimate as denoted by, and the independent variables we will denote with are presented. The summary tables show the parameters obtained by the different methods, such as the certified values (Cv), as well as the number of iterations (it), the calculation time (CPU) and the sum of squared errors (SSE). The assessment was performed on a computer with the following characteristics: Intel® Core™ i7 CPU X990 @ 3.47GHz, 24 Gb of RAM, and a 1 Terabyte hard drive with a 64-bit Windows 7 Professional operating system.

3.1. Problems with a high grade of difficulty

3.1.1. Thurber

$$y = \frac{\beta_1 + \beta_2 x + \beta_3 x^2 + \beta_4 x^3}{1 + \beta_5 x + \beta_6 x^2 + \beta_7 x^3}$$

This problem consists of seven parameters and 37 observations, and corresponds to a study that involves the mobility of a semiconductor electron. The response variable is the measurement of electron mobility, and the independent variable is the natural density logarithm.

3.1.2. BoxBOD

$$y = \beta_1 \left(1 - e^{-\beta_2 x}\right)$$

This problem presents 2 parameters and 6 observations. The response variable is the biochemical demand for oxygen (BOD), and the independent variable is the incubation time in days.

3.1.3. Rat42

$$y = \frac{\beta_1}{1 + e^{\beta_2 - \beta_3 x}}$$

TABLE 1. RESULTS OBTAINED FOR THE THURBER PROBLEM

θ	G-N	G-N Amort.	BFGS	Dogleg	C.V
β_1	1288.1397	1288.1397	1289.4424	1288.1397	1288.1397
β_2	1491.0793	1491.0793	1488.2110	1491.0793	1491.0793
β_3	583.2384	583.2384	580.6049	583.2384	583.2384
β_4	75.4166	75.4166	74.9928	75.4166	75.4166
β_5	0.9663	0.9663	0.96214	0.9663	0.9663
β_6	0.39797	0.39797	0.39859	0.39797	0.39797
β_7	0.04973	0.04973	0.05148	0.04973	0.04973
it	53	100	100	100	
CPU	00:00.5	00:00.8	00:01.1	00:00.9	
SSE	5642.7082	5642.7082	5800.2368	5642.7082	5642.7082

TABLE 2. RESULTS OBTAINED FOR THE BOXBOD PROBLEM

θ	G-N	G-NAmort.	BFGS	DFP	Dogleg	C.V
β_1	213.8094	213.8094	213.8094	213.8094	213.8094	213.8094
β_2	0.54724	0.54724	0.54724	0.54724	0.54724	0.54724
it	14	13	28	28	13	
CPU	00:00.4	00:00.2	00:00.4	00:00.5	00:00.3	
SSE	1168.0089	1168.0089	1168.0089	1168.0089	1168.0089	1168.0089

This problem consists of three parameters and nine observations and corresponds to data from an example of how to adjust sigmoidal growth curves taken from Ratkowsky (1983). The response variable is grass growth, and the independent variable corresponds to time.

3.1.4. Bennett5

$$y = \beta_1 \left(\beta_2 + x\right)^{-\frac{1}{\beta 3}}$$

This problem consists of 3 parameters and 154 observations and corresponds to the magnetization modeling of superconductivity. The response variable is the force of the magnetic field, and the independent variable is the time in minutes.

3.1.5. Eckerle4

$$y= \frac{\beta_1}{\beta_2} e^{-\frac{(x-\beta_3)^2}{2\beta_2^2}}$$

This problem presents 3 parameters and 35 observations and corresponds to the study of the transmittance of circular interference. The response variable is the transmittance, and the independent variable is the wavelength.

3.1.6. Rat43

$$y = \frac{\beta_1}{\left(1 + e^{(\beta_2 - \beta_3 x)}\right)^{\frac{1}{\beta 4}}}$$

This problem consists of 4 parameters and 15 observations and corresponds to a study on how to adjust sigmoidal growth curves (Ratkowsky, 1983). The response variable is the dry weight of onion bulbs and cymes, and the independent variable corresponds to time.

The algorithms implemented present good behavior in general, succeeding in obtaining in most cases the certified values given by the STRD. The BFGS, DFP and L-M methods did not reach the certified values for the Bennett5 model for three parameters and 154 observations, but only for parameter β_1. In one case, it was due to the fact that the maximum number of iterations (DFP) was reached and in the other cases (BFGS and L-M) it was in the presence of numerical instabilities.

TABLE 3. RESULTS OBTAINED FOR THE RAT42 PROBLEM

θ	G-N	G-NAmort.	BFGS	DFP	Dogleg	L-M	C.V
β_1	72.4622	72.4622	72.4622	72.4622	72.4622	72.4622	72.4622
β_2	2.6181	2.6181	2.6181	2.6181	2.6181	2.6181	2.6181
β_3	0.06736	0.06736	0.06736	0.06736	0.06736	0.06736	0.06736
it	7	6	23	27	6	13	
CPU	00:00.3	00:00.4	00:00.4	00:00.6	00:00.3	00:00.5	
SCE	8.05652	8.05652	8.05652	8.05652	8.05652	8.05652	8.05652

TABLE 4. RESULTS OBTAINED FOR THE BENNETT5 PROBLEM

θ	G-N	G-NAmort.	BFGS	DFP	L-M	C.V
β_1	-2523.5058	-2523.5058	-1501.7441	-1502.2288	-1500.0298	-2523.5058
β_2	46.7365	46.7365	41.1951	41.1997	41.1828	46.7366
β_3	0.93218	0.93218	1.0321	1.03204	1.0324	0.93218
it	6	66	26	100	15	
CPU	00:00.5	00:00.8	00:00.9	00:00.7	00:00.6	
SCE	0.000524	0.000524	0.000608	0.000608	0.000609	0.000524

TABLE 5. RESULTS OBTAINED FOR THE ECKERLE4 PROBLEM

θ	G-N	G-NAmort.	BFGS	DFP	Dogleg	L-M	C.V
β_1	1.5544	1.5548	1.5545	1.5542	1.5540	1.5545	1.5544
β_2	4.0889	4.0915	4.0907	4.0882	4.0873	4.0912	4.0888
β_3	451.5411	451.5413	451.5427	451.5387	451.5412	451.5411	451.5412
it	4	3	16	33	7	3	
CPU	00:00.3	00:00.4	00:00.6	00:00.5	00:00.3	00:00.4	
SCE	0.001464	0.001464	0.001464	0.001464	0.001464	0.001464	0.001464

TABLE 6. RESULTS OBTAINED FOR THE RAT43 PROBLEM

θ	G-N	G-NAmort.	BFGS	DFP	Dogleg	L-M	C.V
β_1	699.6415	699.6415	699.6415	699.6415	699.6415	699.6415	699.6415
β_2	5.2771	5.2771	5.2771	5.2771	5.2771	5.2771	5.2771
β_3	0.7596	0.7596	0.7596	0.7596	0.7596	0.7596	0.7596
β_4	1.2792	1.2792	1.2792	1.2792	1.2792	1.2792	1.2792
it	14	100	26	29	39	100	
CPU	00:00.5	00:00.6	00:01.1	00.01.7	00:00.9	00:00.7	
SCE	8786.4049	8786.4049	8786.4049	8786.4049	8786.4049	8786.4049	8786.4049

Figure 1. Adjustments made to Thurber, BoxBOD, Rat42 and Bennett5 models

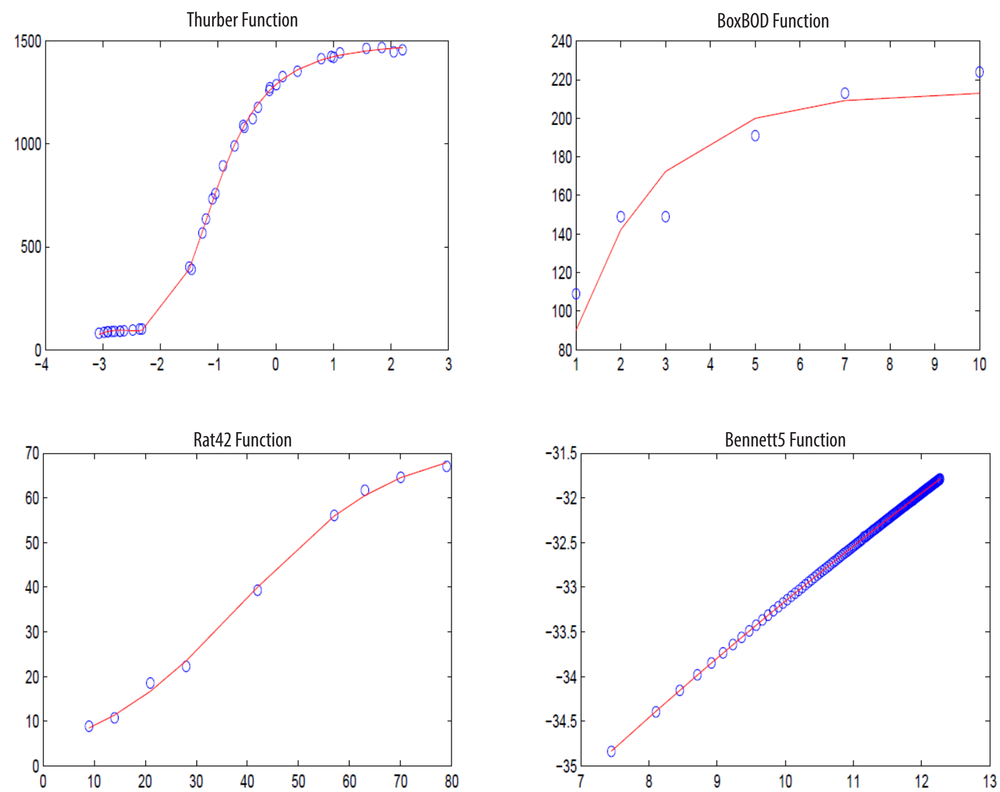
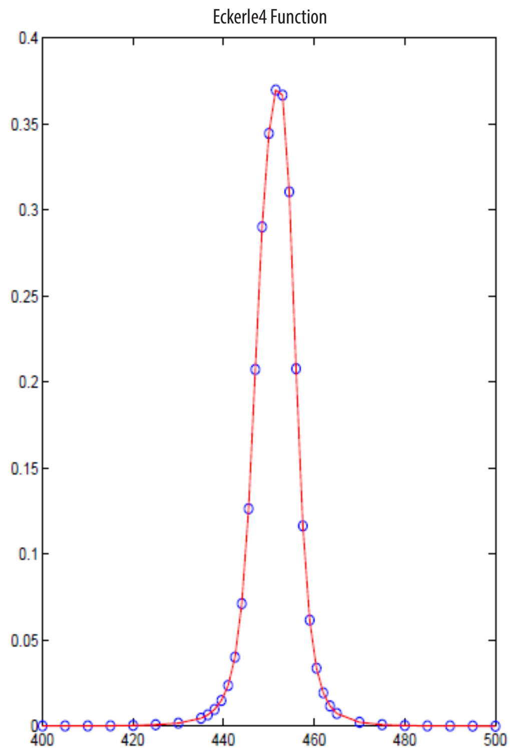


Figure 2. Adjustments made toEckerle4 and Rat43 models



4. APPLICATION IN THE FISHING SECTOR

4.1. Introduction

In Chile fishing is one of the most important industrial sectors in terms of quantities extracted from a natural resource. This is thanks to the great abundance of marine resources Chile has, and it is for that reason that the growth and mortality studies of the different species that make up this important resource are of great interest to those responsible for the sector's exploitation and conservation. Said studies are principally geared towards supporting the knowledge necessary for maintaining balance in exploited stocks in such a way that they become sustainable over time. The results obtained for the growth of sardines, anchovies and hake will be studied and explained. This phenomenon is one of the most important and well-known cases worldwide as an example of nonlinear behavior. The growth of a living being can be divided into two or three very pronounced stages in which it is possible to observe very different growth speeds or rates. In general, living beings experience a first period of growth at very high rates, where a large size is registered in a relatively short time. Later, a second state begins where said rate notably decreases until a certain level of possessing almost a null rate and the size tends to stagnate, see Blasco (N.D.). In this study, it is possible to find concepts related to growth that can be of great interest, such as growth speed, a concept that can help to obtain information regarding the optimum time of interest to take a particular action on a determined living being with the aim of being able to obtain some benefit from it (reproduction time, catch age, sexual maturity, etc.).

In addition, concepts can appear such as growth acceleration, which indicates the absolute variation rate with time, or relative growth rate, which can be useful for comparing trends. However, the present research focuses on the fishing sector with an interest in obtaining the growth parameters of a model with significant applicability in said sector, such as the Von Bertalanffy equation.

4.2. Von Bertalanffy model

In the fishing sector a widely-used growth model is the Von Bertalanffy equation, which represents growth in length and weight, both based on the lifespan of the fish. This is an individual growth model and applies to a large majority of fish. The Von Bertalanffy model offers the allure of trying to deduce the equation based on the anabolism and catabolism rates of the animal, which ideally would be obtained through laboratory experiments (Blasco, 1999). It can be inferred that length is modelled by the equation

$$\hat{L}(t)=L_{\infty}(1-e^{-k(t-t_0)})+\epsilon_t \qquad (22)$$

while growth in weight is estimated through

$$\hat{W}(t)=W_{\infty}(1-e^{-k(t-t_0)})^b+\epsilon_t \qquad (23)$$

where t is the age of the species, $\hat{L}(t)$ and $\hat{W}(t)$ are the average length and weight of the species L_{∞} y W_{∞} over time t, represent the average asymptotic length and weight for the respective species, k is the parameter of the curvature, t_0 is the theoretical age at length zero for a particular species (constant, which represents the age that the fish must supposedly have so its length is equal to zero), b represents the slope of the length-weight relationship, and ϵ_t represents the error. One of the principal applications of the knowledge of the age is, together with the length and weight of the fish, estimation of the growth curve. In addition, it permits the construction of age-size keys which allow the structures by age of catch and stock to be known. Other important parameters that can be estimated are: age of first maturity and spawning; age of recruitment; age of first catch, etc. This makes it possible to estimate abundance, biomass and mortality per fish, fundamental for assigning catch quotas and maintaining the resource. **Figure 3** shows the behavior of the Von Bertalanffy growth model for different values of the parameter k. The importance of this parameter can clearly be seen in a growth model, where a value of k=0 would indicate the maximum length reached for a fish for which the growth factor would be null. When it is not possible to depend on length-age information for individual juveniles, which generates a slant in the distribution of the sizes, the backcalculation technique is used for length to preterit ages. For this, it is necessary to establish a relationship or proportionality between the growth increase of the otolith and the fish, represented by a linear or potential regression. With the resulting expression, it is feasible to calculate the lengths that the fish had when a growth ring was formed. The otoliths constitute a very important part of the internal ear of bone fish. Once the ages are obtained, the procedure used to determine the parameters of the Von Bertalanffy model is with a linearized Beberton and Holt (1957) expression de Beberton y Holt (1957). This method has become one of the cornerstones of fishery biology, given its use as a submodel in more complex models for describing the dynamics of fish populations. The model for weight was only obtained for the sardine because the respective information was not available in all cases.

4.3. Growth of the anchovy (*Engraulis ringens*)

The anchovy is a pelagic fish with a wide geographic distribution in the southeast Pacific. It is distributed from the north of Peru to the X region of Los Lagos, Chile. This species exhibits a short life cycle with three to four years of longevity, reaching a first sexual maturity size of 11.5 cm. Furthermore, it has rapid growth and an elevated natural mortality rate, and it forms dense shoals and as a species is greatly affected by environmental factors in all ages of its life cycle. The anchovy in Chile constitutes a fishery resource that sustains important fishing industries, its destination being primarily fish meal and fish oil production.

4.3.1. Results

In the study conducted on the anchovy the Von Bertalanffy model was applied and the estimation performed based on an observed set of 1,268 specimens, both male and female because the growth of both sexes is similar. As can be observed in **Figure 4**, it shows an almost constant growth curve until approximately 4.5 years, the age at which it reaches its maximum growth size. After this age the anchovy tends to decrease in growth, reaching approximately 16 cm. The data were obtained from the work of Cisterna (2006) and her study estimated the following parameters for the Von Bertalanffy model:

$$\hat{L}(t)=18.47334(1-e^{-0.31685(t+0.71061)})$$

With these parameters, a remainder of R = 840.7997954 was obtained. Upon applying the algorithms shown in **Table 7**, the parameters for the model were estimated. Thus, the following model is obtained

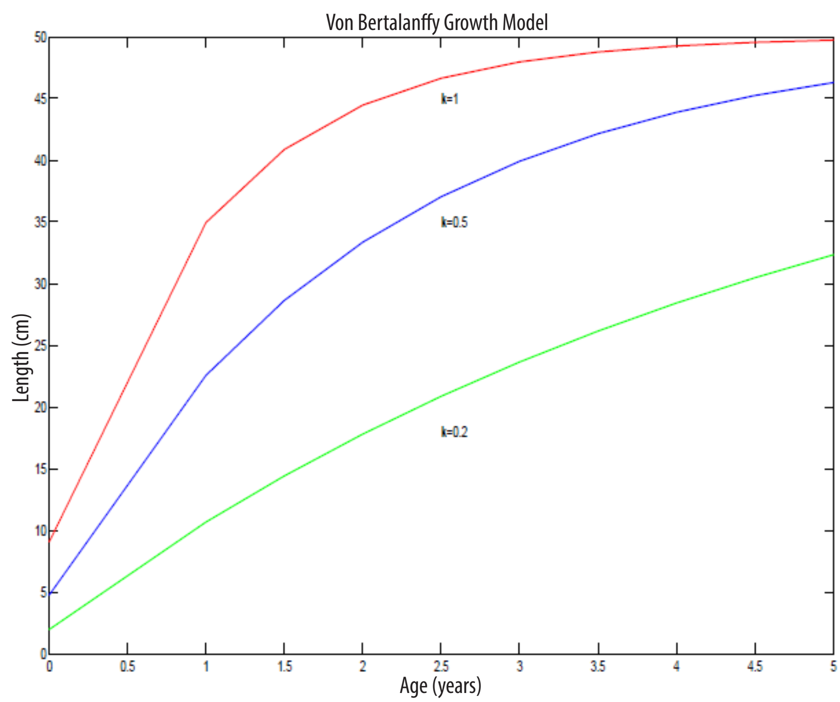
$$\hat{L}(t)=18.42796(1-e^{-0.31636(t+0.70458)})$$

considering the following as initial values: $L_\infty = 19$, $k = 0.5$ año^{-1} $t_0 = -0.5$. The remainder obtained through these parameters was R = 842.6249804, which is smaller than that obtained by Cisterna. **Table 7** summarizes the behavior of the programmed methods as well as the execution times.

4.4. Growth of the common sardine (Strangomera bentinki)

This resource is the second in importance in fishing activity in the central-south region of Chile. For more information on this species. See Peña (2008).

Figure 3. Behavior of Von Bertalanffy for different values of *k*



4.4.1. Results

For the common sardine the Von Bertalanffy model was applied and the estimation was performed for a total of 792 observed samples between female and male because no major difference is distinguishable in the growth of the two sexes. **Figure 5**, as in the case of the anchovy, shows a tendency towards near constant growth, in this case until approximately 3 years. Later, the fish does not exhibit an important growth, reaching 17 cm. The data were obtained from the study conducted by Peña, (2008). Her research estimated the following parameters for the Von Bertalanffy model:

$$\hat{L}(t) = 15.78(1 - e^{-0.686(t+0.1816)})$$

With these parameters, a remainder of R = 1.536. 21968 was obtained. The algorithms presented in this study obtained equal results for all parameters. Thus, the estimated model remains:

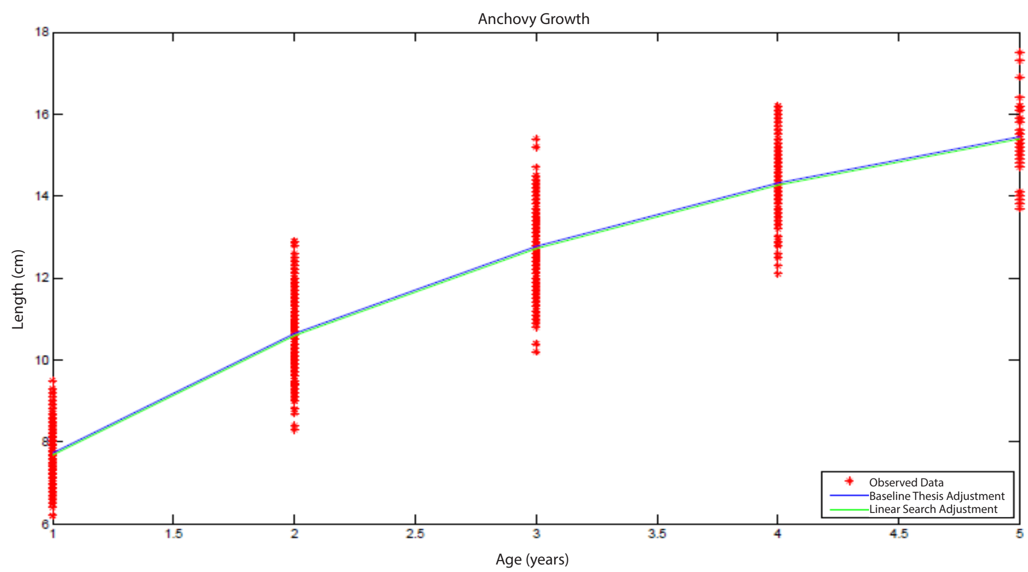
$$\hat{L}(t) = 15.7844(1 - e^{-0.68649(t+0.18162)})$$

The remainder obtained with these parameters was R = 1,536, 18379. The following initial values were used: $L_\infty = 18$, $k = 0.45$ year^{-1}, $t_0 = -0.82$. Below, in **Table 8**, the behavior of the programmed methods is summarized:

TABLE 7. RESULTS OBTAINED BY THE ALGORITHM FOR ANCHOVY GROWTH

θ	G-N	G-NAmort.	BFGS	DFP	Dogleg	L-M
L_∞	18.42796	18.42796	18.42796	18.42796	18.42796	18.42796
k	0.31636	0.31636	0.31636	0.31636	0.31636	0.31636
t_0	-0.70458	-0.70458	-0.70458	-0.70458	-0.70458	-0.70458
it	4	4	37	87	4	9
CPU	00:02.2	00:01.4	0:06.3	00:14.3	00:01.5	00:02.1

Figure 4. Adjustment to anchovy growth



4.5. Growth of the common hake (*Merluccius gayi gayi*)

This resource is one of the most important for fishing activity in the central south region of Chile. For further information about this species: INPESCA, 2007.

4.5.1. Results

4.5.2. Common hake (male)

For the male samples of the common hake 2.019 units were observed and can be found in the next figure below, which shows the length-age relationship. The data were obtained from INPESCA and the model proposed by Neira (2006) estimated the following parameters for the Von Bertalanffy model:

$$L_t = 49.85[1 - e^{-0.44284(t+0.06869)}]$$

With these parameters, a remainder of R = 47.808. 50972 was obtained. Furthermore, all the algorithms programmed in this research project gave the same values for all parameters. We obtain, thus, the following model:

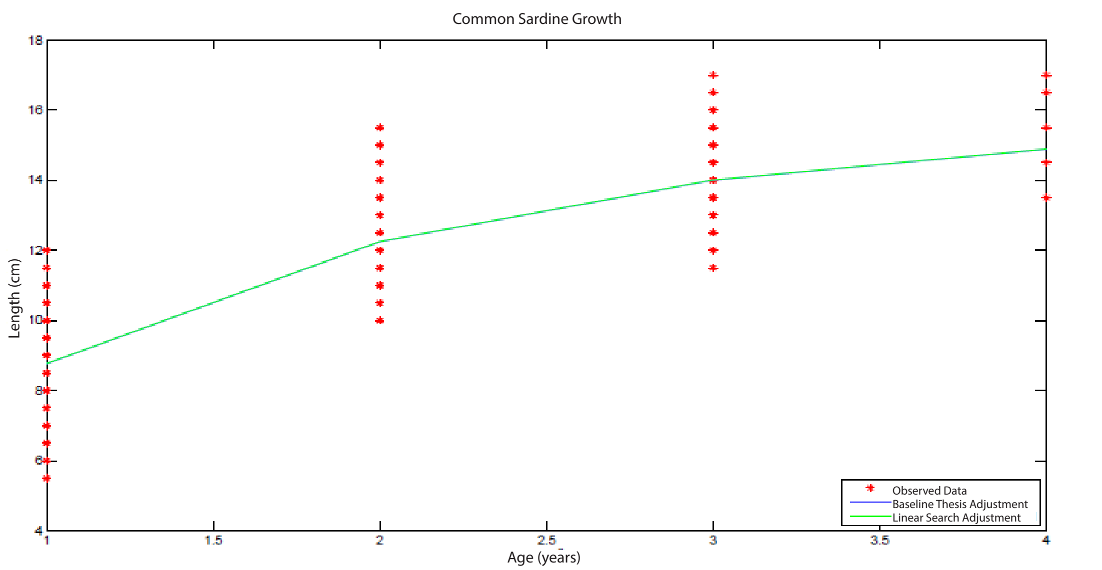
$$L_t = 52.44[1 - e^{-0.36356(t-0.21240)}]$$

with a remainder of R = 16.178. 744. The remainder decreases threefold, which involves an improved adjustment to the experimental data. **Figure 6** shows the data and just one adjustment.

TABLE 8. RESULTS OBTAINED BY THE ALGORITHM FOR COMMON SARDINE GROWTH

θ	G-N	G-NAmort.	BFGS	DFP	L-M
L_∞	15.78437	15.78437	15.78437	15.44615	15.19560
k	0.68649	0.68649	0.68649	0.72485	0.72082
t_0	-0.18162	-0.18162	-0.18162	-0.16229	-0.10822
it	4	22	50	100	100
CPU	00:02.3	00:02.1	00:05.3	00:08.3	00:03.6

Figure 5. Adjustment to common sardine growth



Once the estimation is performed by the least squares method, we note that, of the 6 methods programmed, all garner the same results. Only the BFGS and DFP methods do not converge with each other; however, the moment the process stops they converge at the same parameters as those that did. The results obtained by the algorithm used the following initial values: $L_\infty = 52$, $k = 0.36$ year^{-1}, $t_0 = -0.25$. **Table 9** displays the behavior of the programmed methods.

4.5.3. Common hake (female)

For the female samples of the common hake 3.517 units were observed. They can be found in the next figure below which provides the length-age relationship. As for the males, the data were obtained from INPESCA, and the Von Bertalanffy model estimated by Neira (2006) is the following:

$$L_t = 58.21[1 - e^{-0.28758(t+0.17412)}]$$

obtaining a remainder of $R = 46.709.7978$. All the algorithms presented in this study obtained the same values for the parameters of the Von Bertalanffy model; thus, we obtain the following model:

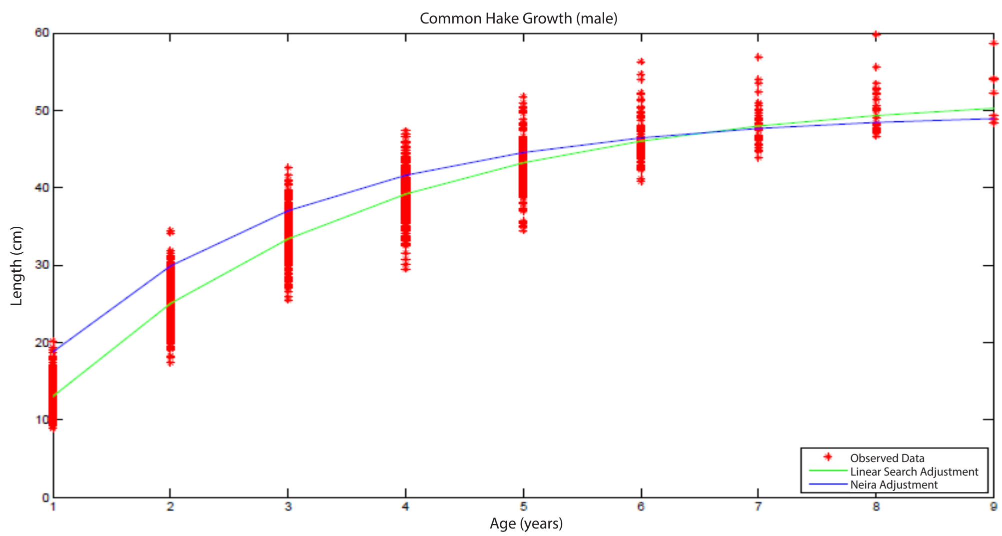
$$L_t = 58.96[1 - e^{-0.29188(t-0.10074)}]$$

with a remainder of $R = 37.860.5497$, utilizing the initial values given by $L_\infty = 69$, $k = 0.19$, $t_0 = -0.65$. The algorithm generates a decrease in the remainder by 1.2 times, involving an improved adjustment to the experimental data. **Figure 7** shows the data and the adjustment with the algorithm. As can be observed, neither of the two curves correctly represent the behavior of the data.

TABLE 9. RESULTS OBTAINED BY THE ALGORITHM FOR COMMON HAKE GROWTH (MALE)

θ	G-N	G-NAmort.	BFGS	DFP	Dogleg	L-M
L_∞	52.43626	52.43626	52.43626	52.43626	52.43626	52.43626
k	0.36356	0.36356	0.36356	0.36356	0.36356	0.36356
t_0	0.21240	0.21240	0.21240	0.21240	0.21240	0.21240
it	6	100	26	31	30	11
CPU	00:02.3	00:16.2	00:09.1	00:10.5	00:05.4	00:02.7

Figure 6. Adjustment to common hake growth (male)



This may be due to the model employed (Von Bertalanffy) not correctly representing the growth of the common hake females for the data in the study. In **Table 10**, the behavior of the programmed methods is shown.

4.6. Cerrato test

Commonly the parameter equality of different sexes can be analyzed by conducting statistical tests that demonstrate that very characteristic. This curve comparison can be performed through a multivariate growth comparison analysis by Cerrato (1990) based on the T^2 Hotelling test. The Cerrato test is a procedure used for growth comparison between the sexes for a determined species. On this occasion the test was applied to the growth of the common hake since the growth for the sardine and anchovy did not present notable differences between the sexes. What occurs with the hake is the female presents a higher development in length than the male. The Cerrato test procedure (1990) can be summarized in the following way: the parameters in comparison are grouped into column vectors defined by

$$\theta_1 = [L_\infty^1\ k_1\ t_0^1]^t \qquad \theta_2 = [L_\infty^1\ k_2\ t_0^2]^t$$

Then the difference between vectors is defined as: $\delta = \theta_1 - \theta_2$. Now it is possible to establish the null and alternative hypotheses respectively:

$$H_0 : \delta = 0$$

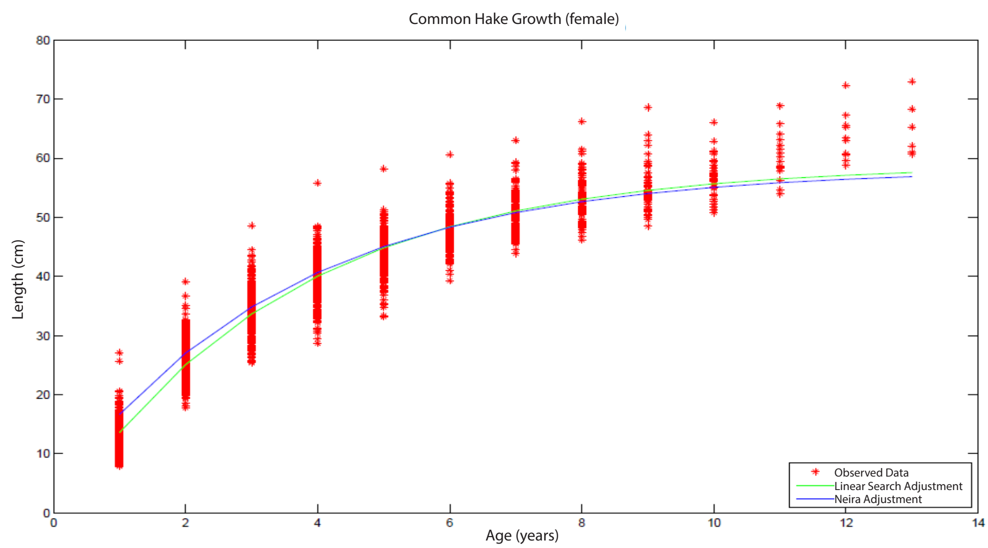
$$H_1 : \delta \neq 0$$

Based on the estimated values, we define: $d = \hat{\theta}_1 - \hat{\theta}_2$. Then, the statistical test is given by the following expression:

TABLE 10. RESULTS OBTAINED BY THE ALGORITHM FOR COMMON HAKE GROWTH (FEMALE)

θ	G-N	G-N Amort.	BFGS	DFP	Dogleg	L-M
L_∞	58.96163	58.96163	58.96163	58.96163	58.96163	58.96163
k	0.29188	0.29188	0.29188	0.29188	0.29188	0.29188
t_0	0.10074	0.10074	0.10074	0.10074	0.10074	0.10074
it	9	100	39	38	100	16
CPU	00:03.2	00:29.1	00:24.4	00:19.2	00:25.1	00:05.2

Figure 7. Adjustment to common hake growth (female)



$$T^2 = d^t U^{-1} d$$

where $U = U_1 + U_2$, and U_i represent the covariance matrix. The expression for matrix U is given by:

$$U = S^2[F^t(\theta)F(\theta)]^{-1}$$

where

$$S^2 = \frac{1}{n-3} S(\hat{\theta})$$

and

$$S(\hat{\theta}) = \sum_{i=1}^{n} (y_i - f(\theta, t_i))^2$$

represent the sum of the squared error. $F(\theta)$ represents the partial derivative matrix with respect to θ_j. Moving on with the procedure, is done $f_i = n_i - 3$, for which f^* is determined by

$$\frac{1}{f^*} = \frac{1}{f_1}\left(\frac{d'U^{-1}U_1U^{-1}d}{d'U^{-1}d}\right)^2 + \frac{1}{f_2}\left(\frac{d'U^{-1}U_2U^{-1}d}{d'U^{-1}d}\right)^2$$

The value of f^* should be among the smallest of the values of f_1 and f_2 and their sum. The null hypothesis is rejected if

$$T^2 > T^2_\alpha(3, f^*)$$

where $T^2_\alpha(p,m)$ is defined as

$$T^2_\alpha(p,m) = \frac{mp}{m-p+1} F^\alpha_{p,m-p+1}$$

with being $F^\alpha_{p,m-p+1}$ the percentage point above the distribution F with p and $m - p + 1$ degrees of freedom. This test can also be used in a bivariable comparison (for example L_∞ and k) or even univariable. The results obtained from the test for the common hake are shown in **Table 11**.

TABLE 11. RESULTS FOR APPLICATION OF THE CERRATO TEST TO THE COMMON HAKE

Species	T^2	$T^2_\alpha(3, f^*)$
Common hake	269.83096982076	14.0380671612343

As can be observed in **Table 11**, it becomes evident that the growth parameters are not the same for both sexes of the common hake, thus it is confirmed that there is a difference in growth between females and males.

5. CONCLUSIONS

The Von Bertalanffy model was applied for the growth of the common hake, anchovy, and common sardine, estimating the parameters of said growth model by way of the application of the distinct parameter estimation algorithms in nonlinear models. The results obtained by the optimization algorithms have been disclosed in their respective sub-applications, meaning for the anchovy, sardine, and common hake. According to them, for the study conducted for the anchovy all the methods succeeded in converging to the values expected according to Cisterna (2008). In the case of the study for the common sardine, the results were different because DFP, like Levenberg-Marquardt, did not completely converge at the values obtained in (2008), although this was due principally to the computational force that these methods demonstrated in the process of determining the size of the step. The Dogleg method doesn't succeed in converging because the algorithm detects singularity in the evaluation of the approximation matrix of the Hessian matrix. In the case of the studies conducted for the common hake the results obtained by the algorithms implemented were significantly different from those presented by Neira (2006). The programmed methods give identical results, but they are different both in the value of the parameters and the remainders. Upon analyzing the parameters obtained for the different species of fish, we can conclude that for the parameter L_∞, it is correct to describe it, in practice, as the average maximum length that could be reached for the different species studied, see [25]. This is to say that for fish like the anchovy and the sardine the values would be $L = 18.42796$ and $L = 15.78437$, respectively. In terms of the parameter k, this indicates the degree of inclination of the growth curve so that for very high values it indicates that the fish has a short lifespan, which is why its growth rate is so elevated. This occurs in the case of the sardine, whose short lifespan makes a value of k = 0.68649 necessary for this parameter. On the other hand, hake (Ojeda & Olivares, 1997) has a longer lifespan, so the value of the parameter k is much lower. Regarding the parameter t_0, it corresponds to a fictitious value associated with a period of the fish in which there is no information (Schnute, 1981). This means it corresponds to the time from when it is born to when it is 1 year old. In this period the fish experiences an exponential growth that stops upon reaching 1 year old (turning point) and begins an exponential decrease. In other words, this parameter corresponds to the age of the fish when it has a length of zero. In general, we can conclude that the algorithms implemented behave well when it comes to the CPU time used to obtain the parameters, the number of iterations and the estimated values of the parameters for highly difficult problems. Regarding application in the fishing sector, the algorithms implemented obtain better results than those reported in the literature (sum of minor squared errors), making them a good option for problem solving in other areas.

REFERENCES

Arancibia, H.; Neira, S. (2006). *Proyección de mediano plazo de la biomasa de merluza común de chile central mediante un enfoque ecotrófico multiespecífico*. Chile: Concepción University.

Beverton, R.J.H.; Holt, S. (1957). *On the Dynamics of Exploited Fish Populations*. London: H.M. Stationery Office. 533 p.

Berghen, F.V. (2004). *Levenberg-Marquardt Algorithms Vs Trust Region Algorithms*. Belgium: Université Libre de Bruxelles. [Online]. Available at: http://www.applied-mathematics.net/LMvsTR/LMvsTR.pdf

Blasco, A. (N.D.). *La descripción del crecimiento*. Informe técnico ocasional Nº6. Valencia, España: Departamento de Ciencia animal, Universidad Politécnica de Valencia.

Bonnans, J.F.; Gilbert, J.C.; Lemarechal, C.; Sagastizábal, C.A. (2002). *Numerical Optimization: Theoretical and Practical Aspects.* New York, U.S.A.: Springer-Verlag New York, Inc.

Börlin, N. (2007). *Nonlinear Optimization: trust-region methods*. Umeå, Sweden: Department of Computing Science, Umeå University. [Online]. Available at: https://www8.cs.umu.se/kurser/5DA001/HT07/lectures/trust-handouts.pdf

Cerrato, R.M. (2011). Interpretable Statistical Tests for Growth Comparisons Using Parameters in the Von Bertalanffy Equation. *Canadian Journal of Fisheries and Aquatic Sciences*, 47(7), April, pp. 1416-1426.

Cisterna A., L. E. (2008). *Análisis del Crecimiento y Estimación de la Mortalidad natural de "Engraulis ringens"(Jenyns, 1842) (Clupeiformes, engraulidae) en la zona centro sur de Chile, durante el año 2006*, tesis (Biología Marina), Chile, Facultad de Ciencias, U.C.S.C.

Coleman, T.F. (1984). *Large Sparse Numerical Optimization*. Springer-Verlag Berlin Heidelberg.

De la Fuente, J.L. (1995). *Tecnologías computacionales para sistemas de ecuaciones, optimización lineal y entera*. Editorial Reverte S.A. 762 p.

Draper, N.R.; Smith, H. (1966). *Applied Regression Analysis*. U.S.A. John Wiley & Sons.

Fletcher, R. (1980). *Practical Methods of Optimization.* U.S.A. John Wiley & Sons.

Frandsen, P.E.; Jonasson, K.; Nielsen, H.B.; Tingleff, O. (2004). *Unconstrained Optimization*. Denmark: Technical University of Denmark. [Online]. Available at: file:///Users/juanignacioplasenzotti/Downloads/imm3217%20(1).pdf

Kelley, C.T. (1995). *Frontiers in Applied Mathematics: Iterative Methods for Linear and Nonlinear Equations.* North Carolina, U.S.A.: North Carolina State University, Society for Industrial and Applied Mathematics.

Lange, K. (2004) *Optimization*. U.S.A. Springer-Verlag New York, Inc.

Luenberger, D.G. (1984) *Linear and Nonlinear Programming*. Addison-Wesley Publishing Company.

Madsen, K.; Nielsen, H.B.; Tingleff, O. (2004). *Methods for Non-linear Least Squares Problems*. Denmark: Technical University of Denmark. [Online]. Available at: http://www.ltu.se/cms_fs/1.51590!/nonlinear_least_squares.pdf

Mizutani, E. (1999). *Powell's Dogleg Trust-Region Steps with the Quasi-Newton Augmen- Ted Hessian for Neural Nonlinear Least-Squares Learning*. Berkeley, CA, U.S.A.: Institute of Electrical and Electronics Engineers.

Neter, J.; Kutner, M.; Nachtsheim, C.; Wasserman W. (1996). *Applied Linear Regression Models*. 4th edition. U.S.A.: McGraw-Hill/Irwin.

Nocedal, J.; Wright, S.J. (1999). *Numerical Optimization*. Springer-Verlag New York, Inc.

Ojeda, V.; Olivares, J. (1997). *Validación de los métodos aplicados en la estimación de edad y crecimiento y determinación de la mortalidad en Merluza Común en la zona centro-sur*. Chile: Informes Técnicos del Fondo de Investigación Pesquera - Proyecto #95-15.

Peña-Saez, M.E. (2008). *Análisis de Crecimiento y Estimación de Mortalidad natural de la "Sardina Común" (Strangomera bentincki, NORMAN 1936) en la zona centro sur de Chile*, tesis (Biología Marina), Chile, Facultad de Ciencias, U.C.S.C.

Ratkowsky, D.A. (1983). *Nonlinear Regression Modeling: A Unified Practical Approach*. New york, U.S.A.: Marcel Dekker, Inc. 288 p.

Schnute, J. (1981). A Versatile Growth Model with Statistically Stable Parameters. *Canadian Journal of Fisheries and Aquatic Sciences*, 38(9), April, pp. 1128-1140.

Subsecretaría de Pesca (2007). *Merluza Común, Merluccius gayi gayi (Guichenot, 1848)*. Chile: Subsecretaría de Pesca - Gobierno de Chile. [Online]. Available at: http://www.fip.cl/documentacion/principalespesquerias/archivo_6.pdf

TO REFERENCE THIS ARTICLE / PARA CITAR ESTE ARTÍCULO / PARA CITAR ESTE ARTIGO /

Cornejo-Zúñiga, Ó.; Rebolledo-Vega, R. (2016). Estimation of Parameters in Nonlinear Models: Algorithms and Applications. *Revista EIA*, 13(25), January-June, pp. 81-98. [Online]. Available at: DOI: http:/dx.doi.org/10.14508/reia.2016.13.25.81-98