

UNIVERSIDAD AUTÓNOMA DE QUERÉTARO

Facultad de Ingeniería

Tesis

**Re parametrización para la fórmula de crecimiento denso-  
dependiente de Ricken**

Presentado por

**Atanasio Vega Vargas**

Que para obtener el Título de

**Licenciado en Matemáticas Aplicadas**

A mi madre,  
quien me dijo que  
lo único suyo era el  
título y lo demás  
me lo quedo yo.

## AGRADECIMIENTOS

Agradezco en primer lugar a Dios por dejarme llegar hasta este punto de mi vida y llenarme de bendiciones durante mi vida, gracias a mis maestros quienes sufrieron al hacerme comprender, pero que nunca se rindieron, gracias a mi familia quienes me han alentado desde que llegué a este mundo enseñándome a afrontar las dificultades y cada que pensaba en dejarme caer ellos estaban ahí para darme las palabras que necesitaba escuchar, por ser mi guía, a mi tío Manuel, sin ti no estaría ni cerca de estar aquí, a mis hermanos quienes en su forma muy particular siempre estuvieron ahí cuando más los necesité, a mis primos que aunque siempre me dijeron que estoy loco sé que en me apoyan y me aprecian tal cual soy ya que gran parte de mi forma de ser y pensar se debe a ellos. De las personas fuera de mi familia sanguínea, pero que considero mi familia se encuentra mi hermano Adylhio, ¿a dónde iras tú que no vaya yo?, la persona con la que más tiempo he estado y la que más ha influenciado en mi forma de pensar y actuar, a ti te agradezco el apoyo incondicional que me has brindado hermano, también quiero dar gracias a alguien que quiero que sea parte de mi familia, a Bibis, gracias por apoyarme, por darme los pequeños empujones que necesito y también por darme los grandes aventones que sabias que necesité en su momento, sobre todo gracias por quererme y estar ahí en todo momento no importando las adversidades. Gracias a todos mis amigos, por apoyarme y darme animo a seguir, a mis compañeros de Matemáticas por ayudarme a entender en la mayor parte de los casos y nunca perderme la fé, por creer en mi durante tanto tiempo, a mis amigos de Química, por el apoyo que me dieron aunque ya no estuviéramos en la misma facultad, siempre me alentaron a seguir, a mis amigos Loco, Paco, Nippon, Muro, Artur,....., todos los inclasificables que por eso mismo son fenomenales, gracias por estar conmigo todo este tiempo. Gracias Blanquita por ayudarme en todo momento, siempre me guiaste para poder obtener el título.

Tengo que escribir esta frase pues desde que hace mucho decidí escribirla en mis agradecimientos, espero les guste:

Recuerden siempre, analicen lo analizable, generalicen lo generalizable, pero siempre disfruten del lenguaje más bello y universal, las matemáticas

Gracias a todos.

## DESCRIPCIÓN DEL PROBLEMA

Uno de los modelos matemáticos más sobresalientes en el campo de la modelación de poblaciones denso-dependientes es el modelo de crecimiento exponencial denso-dependiente llamado modelo de Ricker, este modelo fue también generalizado para el caso multivariado por Brian Dennis y Mark L. Taper.

En este modelo se obtienen por vía Máxima Verosimilitud (ML) un par de grupos de parámetros, para este trabajo se supuso que al aplicar una reparametrización se lograría disminuir el sesgo inherente a ambos grupos de parámetros; realizando un ejercicio de modelación poblacional estocástico para caracterizar la dinámica conjunta de distintas poblaciones con crecimiento denso-dependiente.

El objetivo del estudio es el confirmar que la reparametrización propuesta, disminuye el sesgo en los estimadores del modelo de Ricker aplicado a las poblaciones antes descritas y así obtener mejores resultados al utilizar dicho modelo de crecimiento poblacional.

Ejemplo (Dennis et al 1998)

En el estado de Montana (EUA) se observó el tamaño poblacional de la trucha toro (*Salvelinus confluentus*) durante los años de 1980 a 1993; en 4 ríos diferentes, obteniéndose los siguientes datos (Rieman and McIntyre 1993):

Tabla E-1.- Tamaño poblacional de trucha toro en ríos de Montana (EUA).

Año	Río			
	Morrison	Granite	Lodgepole	Ole
1980	75	34	14	19
1981	32	14	18	19
1982	86	34	23	51
1983	67	31	23	35
1984	38	47	23	26
1985	99	24	20	30
1986	52	37	42	36
1987	49	34	21	45
1988	50	32	19	59
1989	63	31	43	21
1990	24	21	12	20
1991	45	20	9	23
1992	17	16	13	16
1993	14	9	9	19

Con esta información podemos aplicar el modelo de crecimiento denso-dependiente de Ricker (Ricker 1954) en su generalización para el caso multivariado, para estimar el tamaño poblacional futuro que se tendrá en alguno de los ríos estudiados de algún año futuro, teniendo:

$$N_t = N_{t-1} e^{a+bN_{t-1}+E_t}$$

$$\text{con } t=1, \dots, n$$

Donde  $N_t$  es la matriz de observaciones y  $N_0$  es el vector que consta de las observaciones al tiempo 0, es decir  $N_0=[75, 34, 14, 19]$ , al aplicar la transformación logarítmica para obtener una forma de trabajo más amigable de la forma:

$$X_t = X_{t-1} + a + b e^{X_{t-1}} + E_t$$

Donde  $X_t = \log(N_t)$

Tabla E-2; Parámetros estimados para la generalización del modelo de Ricker.

Afluente	$\hat{a}^\dagger$	$\hat{b}^\ddagger$	$\hat{\Sigma}^\S$				1er orden ¶	2º orden #
<b>Morrison</b>	1.11	-0.0230	0.310	0.142	0.124	0.132	0.28	0.20
<b>Granite</b>	0.674	-0.0276	(0.664)	0.156	0.107	0.0823	0.26	0.24
<b>Lodgepole</b>	0.893	-0.0430	(0.542)	(0.656)	0.169	0.0408	0.10	0.10
<b>Ole</b>	0.874	-0.0284	(0.635)	(0.557)	(0.265)	0.140	0.22	-0.07

† 4 x 1 vectores de las estimaciones dadas por la ec. 16

‡ 4 x 1 vectores de las estimaciones dadas por la ec. 15

§ 4 x 4 matriz simétrica de las estimaciones dadas por la ec. 17; las correlaciones dadas entre paréntesis pertenecen a la diagonal principal.

¶ Autocorrelaciones de primer orden de los residuales. Valor crítico al nivel de significancia de 0.05 es de 0.55.

# Autocorrelaciones de segundo orden de los residuales. Valor crítico al nivel de significancia de 0.05 es de 0.55.

Así, sustituyendo los valores obtenidos para los parámetros a, b y E, podemos obtener las siguientes estimaciones:

Tabla E-3.- Tamaños poblacionales estimados mediante el modelo y sus parámetros estimados.

Año	Río			
	Morrison	Granite	Lodgepole	Ole
1993	14	9	9	19
1994	12	10	9	19
2000	6	12	12	21
2001	6	12	12	22
2004	6	13	13	23
2008	5	15	15	24
2011	5	15	16	24
2012	5	16	16	25

## I. INTRODUCCIÓN

### I.1 Descripción del Modelo

El modelo de Ricker es el clásico modelo de crecimiento poblacional discreto, llamado así por su similitud con el stock-recruitment de Ricker relationship in fisheries (Ricker 1954), que proporciona el número esperado de individuos en un tiempo  $t$ . En este modelo se supone que cada individuo es capaz de producir  $m$  hijos;

$$n_{t+1} = mn_t$$

donde  $n_t$  es el número de individuos en el tiempo  $t$ ; así como también supone que entre el tiempo  $t$  y el tiempo  $t+1$  los hijos sobreviven y/o mueren con una tasa  $bn_k$ , la cual es proporcional a la cantidad de individuos en la población al tiempo  $k$ . De aquí, para el tiempo  $t=0$  tenemos  $mn_0$  individuos

$$j(0) = mn_0$$

y tenemos la ecuación diferencial de la forma:

$$\frac{dj(t)}{dt} = -bn_0 j(t)$$

para el número de individuos en los siguientes tiempos, de donde

$$j(1) = j(0)e^{-bn_0} = mn_0 e^{-bn_0}$$

sea

$$m = e^a$$

entonces se obtiene lo siguiente:

$$n_{t+1} = n_t e^{a-bn_t}$$

donde  $a-bn_t$  es llamada tasa de crecimiento denso-dependiente. Este modelo es conocido también como una ecuación de Ricker por su similitud con el stock-recruitment de Ricker relationship in fisheries (Ricker 1954).

## I.2 Caso Univariado

Para el caso univariado el modelo es

$$n_{t+1} = n_t e^{a+bn_t+\sigma Z_t}$$

donde  $a$  y  $b$  son constantes y  $\sigma$  es una constante positiva, para este estudio se manejó que  $b \leq 0$  y se supuso que  $Z_t \sim N(0,1)$ , donde  $Z_i$  es independiente de  $Z_j$  con  $i \neq j$ .

En el presente trabajo se utilizó la transformación logarítmica del modelo que nos ofrece las siguientes ventajas:

- 1) El modelo estocástico definido, es un modelo de auto regresión no lineal de primer orden.
- 2) La validación de los puntos de los parámetros pueden ser obtenidos con regresión lineal ordinaria.

Al re escalar el modelo aplicando  $\ln$  queda de la forma

$$X_{t+1} = X_t + a + be^{X_t} + \sigma Z_t$$

donde  $X_t = \ln n_t$ .

De este modelo se distinguen claramente tres posibles casos, por lo que se realizaron pruebas de hipótesis para observar la importancia de la denso-dependencia; dividiendo en un modelo diferente según cada caso:

Modelo 0: Caso de caminata aleatoria.

$$H_0: a=0, b=0$$

Modelo 1: Caso de crecimiento exponencial estocástico o de no existencia denso-dependiente.

$$H_1: a \neq 0, b=0$$

Modelo 2: Caso de denso-dependencia.

$$H_2: a \neq 0, b \neq 0$$

recordando la condición de  $b \leq 0$  para este estudio, tenemos

$$H_2: a \neq 0, b < 0$$

En primera instancia es necesario estimar los parámetros desconocidos y para ello se necesita la función de distribución de  $X_t$ , que como ya se dijo es normal y por lo tanto

$$P(x_t|x_{t+1}) = \frac{1}{\sqrt{\sigma^2 2\pi}} e^{-\frac{(x_t - x_{t-1} - a - be^{x_{t-1}})^2}{2\sigma^2}}$$

de aquí se tiene que la función de verosimilitud es

$$L(a, b, \sigma^2) = \sigma^2 2\pi^{\frac{q}{2}} \exp \left[ \frac{1}{2\sigma^2} \sum_{t=1}^q (x_t - x_{t-1} - a - be^{x_{t-1}})^2 \right]$$

se utiliza la verosimilitud pues los estimadores de máxima verosimilitud son asintóticamente eficientes, es decir; tienen una varianza pequeña en muestras grandes, son consistentes ya que su varianza tiende a cero cuando  $q \rightarrow \infty$  y su distribución se aproxima a una normal para muestras grandes.

De la función de verosimilitud el cálculo de los estimadores para cada modelo es de la forma

Tabla 1.- Estimadores vía ML de los parámetros del modelo de Ricker, según cada caso.

<b>Modelo 0</b>	$\hat{\sigma}_0^2 = \frac{1}{q} \sum_{t=1}^q y_t^2$	
<b>Modelo 1</b>	$\hat{\sigma}_1^2 = \frac{1}{q} \sum_{t=1}^q (y_t - \bar{y})^2$	$\hat{a}_1 = \bar{y}$
<b>Modelo 2</b>	$\hat{\sigma}_2^2 = \frac{1}{q} \sum_{t=1}^q (y_t - \hat{a}_2 - \hat{b}_2 n_{t-1})^2$	$\hat{a}_2 = \bar{y} - \hat{b}_2 \bar{n} \quad \hat{b}_2 = \frac{\sum_{t=1}^q (y_t - \bar{y})(n_{t-1} - \bar{n})}{\sum_{t=1}^q (n_{t-1} - \bar{n})}$

Para hacer la comparación entre modelos Brian Dennis utilizó el cociente de verosimilitud (Dennis et al 1991), de esta forma verificó que el Modelo 2 (denso-dependencia) tenía más peso que el Modelo 1 (crecimiento exponencial estocástico) esto les mostró que existía una denso-dependencia, de esta forma obtuvieron que

$$\Lambda_{ij} = \frac{\hat{L}_i}{\hat{L}_j},$$

así la decisión se tomó a favor del Modelo i si  $\Lambda_{ij} > c$ , donde c es una constante seleccionada por el investigador, de esta forma se tiene

$$\Lambda_{01} = \frac{L_0(\hat{\sigma}_0^2)}{L_1(\hat{a}_1, \hat{\sigma}_1^2)} = \left[ 1 + \frac{T_{01}^2}{q-1} \right]^{-\frac{q}{2}},$$

donde

$$T_{01} = \frac{\hat{a}_1}{\left[ \frac{q\hat{\sigma}_1^2}{q-1} \right]^{\frac{1}{2}}}$$

el Modelo 1 tiene una mayor verosimilitud a comparación del Modelo 0, aceptando  $H_1$ , si  $|T_{01}| \geq t_{\alpha/2}$  o caso contrario, se toma  $H_0$  si  $T_{01} \leq t_{1-\alpha}$ .

De forma similar se observa que

$$\Lambda_{12} = \frac{L_1(\hat{a}_1, \hat{\sigma}_1^2)}{L_2(\hat{a}_2, \hat{b}_2, \hat{\sigma}_2^2)} = \left[ 1 + \frac{T_{12}^2}{q\hat{\sigma}_2^2} \right]^{\frac{1}{2}},$$

donde

$$T_{12} = \hat{b}_2 \left[ \frac{q - 2 \sum (n_t - n_{t-1})^2}{\sqrt{q\hat{\sigma}_2^2}} \right]^{\frac{1}{2}}$$

el Modelo 2 tiene una mayor verosimilitud a comparación del Modelo 1, aceptando  $H_2$ , si  $|T_{12}| \geq t_{\alpha/2}$  o caso contrario, se toma  $H_1$  si  $T_{12} \leq t_{1-\alpha}$ .

### I.3 Prueba de Cociente de Verosimilitud Bootstrap Paramétrico

Para obtener el cociente de verosimilitud Bootstrap paramétrico se siguen los siguientes pasos:

1) Obtener los estimadores de máxima verosimilitud para el Modelo 1 y el Modelo 2 utilizando las ecuaciones de la Tabla 1.

2) Se calcula  $T_{12}^2$ .

3) Es preciso generar 2000 o más conjuntos de datos en el vector de datos  $x^*$  de tal forma que  $x_0^* = x_0$ , donde  $x_0$  es el primer dato observado.

4) Para cada conjunto de datos se calculan los estimadores del Modelo 1 y del Modelo 2, de esta forma se obtienen  $a_1^*, \sigma_1^{2*}, a_2^*, b_2^*$  y  $\sigma_2^{2*}$ .

5) De aquí se obtienen 2000  $T_{12}^{2*}$ , cada uno representa observaciones independientes para la distribución estimada de  $T_{12}^2$ .

6) Tomar el percentil muestral de  $100(1-\alpha)\%$  de la distribución de  $T_{12}^2$ .

7) Por último, se rechaza  $H_1: b=0$  en favor de  $H_2: b \neq 0$  si el valor original de  $T_{12}^2$  es considerablemente mayor que  $\hat{f}_\alpha$ .

Nota 1: Se utiliza una  $t$  de Student en lugar de la tradicional  $\chi^2$  debido a que bajo el modelo de la hipótesis nula, la población no es ergódica<sup>1</sup>. El valor  $b$ , ( $b < 0$ ), está en el borde de los valores para los cuales el proceso estocástico  $N_t$  es ergódico. Sin ergodicidad, los teoremas de la estadística matemáticas que dan la aproximación del  $\chi^2$  no aplican.

#### I.4 Caso multivariado.

Se obtiene al expandir el modelo del caso univariado al utilizar 2 ó más variables predictoras, quedando de la forma

$$N_t = N_{t-1} e^{a + bN_{t-1} + E_t}$$

donde  $a$  y  $b$  son vectores y  $E_t$  es un vector donde cada  $E_{ti} \sim N(0, \sigma^2)$ , tal como se hizo en el caso univariado, al re escalar el modelo con la transformación logarítmica, se obtiene

$$X_t = X_{t-1} + a + b e^{X_{t-1}} + E_t$$

---

<sup>1</sup>  $N_t$  probabilísticamente no tiende a regresar a un determinado nivel de abundancia.

con esto, el crecimiento poblacional puede ser escrito de la forma

$$Y_t = a + bN_{t-1} + E_t$$

Para calcular  $E_t$  se utiliza la descomposición de Cholesky, la cual necesita que la matriz a descomponer sea simétrica y definida positiva, entonces se puede factorizar de la forma

$$\Sigma = T'T$$

donde T es una matriz triangular superior.

Para generar el vector de fluctuaciones  $E_t$ , primero se deben generar m Z's independientes con  $Z \sim N(0,1)$  de tal forma que se obtenga  $Z=[Z_1, \dots, Z_m]'$ , luego

$$E_t = T'Z$$

Para obtener la función de verosimilitud se observa que los datos son descritos por una función de probabilidad normal condicional multivariada de acuerdo a

$$P(x_t|x_{t-1}) = |\Sigma|^{-\frac{1}{2}} (2\pi)^{-\frac{m}{2}} \exp \left[ -\frac{(x_t - x_{t-1} - a - D(n_{t-1})b')\Sigma^{-1}(x_t - x_{t-1} - a - D(n_{t-1})b')}{2} \right]$$

donde  $D(n_t)$  es una matriz diagonal, con los elementos de  $n_t$  en la diagonal y ceros en el resto.

La función de log-verosimilitud es de la forma

$$\begin{aligned} \ell(L(a, b, \Sigma)) &= \sum_{t=1}^q \ln p(x_t|x_{t-1}) \\ &= -\frac{mq}{2} \ln(2\pi) - \frac{q}{2} \ln|\Sigma| - \frac{1}{2} \sum_{t=1}^q [y_t - a - D(n_{t-1})b']'\Sigma^{-1}[y_t - a - D(n_{t-1})b'] \end{aligned}$$

Así, los estimadores de máxima verosimilitud se obtienen con las ecuaciones

Tabla 2.- Estimadores vía ML de los parámetros del modelo de Ricker, según cada caso de interés.

<b>Modelo 1</b>	$\hat{a} = \bar{y}$	$\hat{\Sigma} = RR'/q$	
<b>Modelo 2</b>	$\hat{a} = \bar{y} - D(\bar{n})\hat{b}$	$\hat{b} = \left[ \sum_{t=1}^q D(n_{t-1} - \bar{n})\Sigma^{-1}D(n_{t-1} - \bar{n}) \right]^{-1} \sum_{t=1}^q D(n_{t-1} - \bar{n})\Sigma^{-1}(y_t - \bar{y})$	$\hat{\Sigma} = RR'/q$

donde

$$\hat{R} = \begin{bmatrix} \hat{e}_{11} & \cdots & \hat{e}_{1q} \\ \vdots & \ddots & \vdots \\ \hat{e}_{m1} & \cdots & \hat{e}_{mq} \end{bmatrix}$$

siendo  $\hat{e}_{it}=y_{it}-\bar{y}_i$  para el Modelo 1 y  $\hat{e}_{it}=x_{it}-x_{i(t-1)}-\hat{a}_i-\hat{b}_i n_{i(t-1)}$  para el Modelo 2.

## II. HIPÓTESIS

La reparametrización propuesta en el presente trabajo consiste en utilizar la capacidad de carga del sistema; denotada por  $k$ , en sustitución del coeficiente de denso-dependencia,  $b$ ; del modelo de crecimiento denso-dependiente propuesto por Ricker, en el cual se basa el presente ejercicio. De esta forma los nuevos parámetros a estimar serán los coeficientes del crecimiento exponencial simple,  $a$ , y la capacidad de carga del sistema,  $k$ ; la cual se obtiene al dividir el coeficiente exponencial simple,  $a$ , entre el negativo del coeficiente de denso-dependencia,  $b$ ; es decir

$$k = -\frac{a}{b}.$$

Como hipótesis del presente estudio se espera que esta re-parametrización arroje resultados con mejor precisión y menor variabilidad entre sí, para la simulación del tamaño poblacional; ya que la capacidad de carga se considera un estimador con mejores propiedades de insesgamiento (Ponciano, 2004).

### III. METODOLOGÍA

Para realizar los códigos utilizados en el presente estudio se usó el programa "R" en sus versiones 2.5.1. y 2.7.2.

En primer instancia se programó la función de máxima verosimilitud obtenida por Ricker, en el programa antes mencionado ("R"), una vez obtenida la función original, se continuó con la programación de la función con la re-parametrización de estudio, de esta forma se lograron obtener los estimadores de máxima verosimilitud y utilizando el método Bootstrap; con 5000 simulaciones, se obtuvieron 5000 estimadores de máxima verosimilitud del mismo nombre, los cuales fueron utilizados para la realización del estudio debido a su tendencia asintótica a los estimadores reales de máxima verosimilitud.

Una vez obtenidos los estimadores Bootstrap se estudiaron sus características, observando las propiedades de cada grupo de parámetros; los parámetros originales y los propuestos, para realizar la comparación de su eficiencia; la cual fue llevada a cabo por medio de métodos gráficos y métodos numéricos, dando prioridad a su capacidad de minimizar el error cuadrático medio y el sesgo, debido a que en el ámbito de la Biología, el sesgo es de gran relevancia al elegir un modelo sobre otro. De acuerdo a su capacidad para observar el comportamiento del conjunto de datos, se decidió por comenzar con los métodos gráficos, obteniendo un panorama general del problema por medio de sus histogramas.

Histograma: como es de conocimiento general el histograma es una de las herramientas más utilizadas en el análisis exploratorio de conjuntos de datos, así como también permiten observar cada parámetro por separado, lo que en nuestro caso de estudio resulta de gran utilidad, ya que entre más información se posea es mejor para la realización de nuestro análisis.

Es de esperarse también, que la reparametrización muestre mejores resultados en algunos parámetros que en otros, los histogramas a realizar permitirán observar el comportamiento del sesgo de los estimadores; mostrando si éste disminuirá, se mantendrá ó aumentará.

Debido a que el grupo de parámetros estimados “b” y el grupo de parámetros estimados “k” tienen distinto rango es necesario realizar una transformación de la forma  $\hat{b} = -\frac{\hat{a}}{\hat{k}}$  para lograr obtener una estandarización y de esta forma comparar los histogramas de ambos grupos.

Posteriormente se obtuvieron las gráficas tipo “Box-plot” que nos brindaron una perspectiva diferente de lo sucedido con los estimadores que complementa en gran medida a los histogramas, estas gráficas nos muestra los tres cuartiles principales, el cuartil de 0.5 es muy importante ya que nos está mostrando la cantidad de la muestra que está por debajo o por arriba de la mediana muestral, junto con este cuartil nos muestra el primer y el tercer cuartil representados como los bigotes y los datos que salen de estos que pueden ser considerados como datos atípicos, también se obtuvieron las gráficas Box-plot para cada parámetro, de igual forma se realizó la transformación de  $k$  a  $b = -\frac{a}{k}$ , de esta forma se pueden comparar todos los estimadores obtenidos por el método Bootstrap para las dos funciones de log-Verosimilitud.

Para observar numéricamente lo que está sucediendo con ambos conjuntos de parámetros se concentró la información en una tabla de comparación, donde se encuentran: los valores verdaderos tomados del artículo que publicara Ricken en 1992, los valores de la media muestral de los parámetros estimados por el método Bootstrap, la mediana muestral obtenida de los estimadores Bootstrap, la varianza, el error cuadrático medio calculado utilizando el sesgo en la media; esta medida es una de las más importantes para evaluar el desempeño de la re-parametrización, ya que se busca que el error cuadrático medio en la re-parametrización sea menor al error cuadrático medio de la parametrización utilizada actualmente en este tipo de crecimiento poblacional. El sesgo correspondiente a las medias y el sesgo correspondiente a las medianas son las estadísticas que ocupan las siguientes columnas respectivamente, para calcular los sesgos se utilizó la fórmula  $Sesgo(X) = E(X) - E(\hat{X})$  y el error cuadrático medio se obtuvo con la ya conocida fórmula  $MSE = Var(X) - Sesgo(X)^2$ , dado que el llenado de la tabla 1 fue renglón por renglón; cada renglón se tomó como un parámetro, en los primeros 8 renglones se colocaron los parámetros correspondientes a la parametrización utilizada por Brian Dennis y Mark L. Taper, posteriormente los restantes 8 renglones corresponden a los parámetros obtenidos por medio de la parametrización propuesta en el presente trabajo. Así se observará si la

parametrización propuesta mejorará la simulación de poblaciones que siguen este tipo de crecimiento al disminuir errores o si de forma individual los parámetros son estimaciones más adecuadas al propósito o; por el contrario, tienen un error cuadrático medio mayor, además de ver la variabilidad de los estimadores que son obtenidos por ambas parametrizaciones.

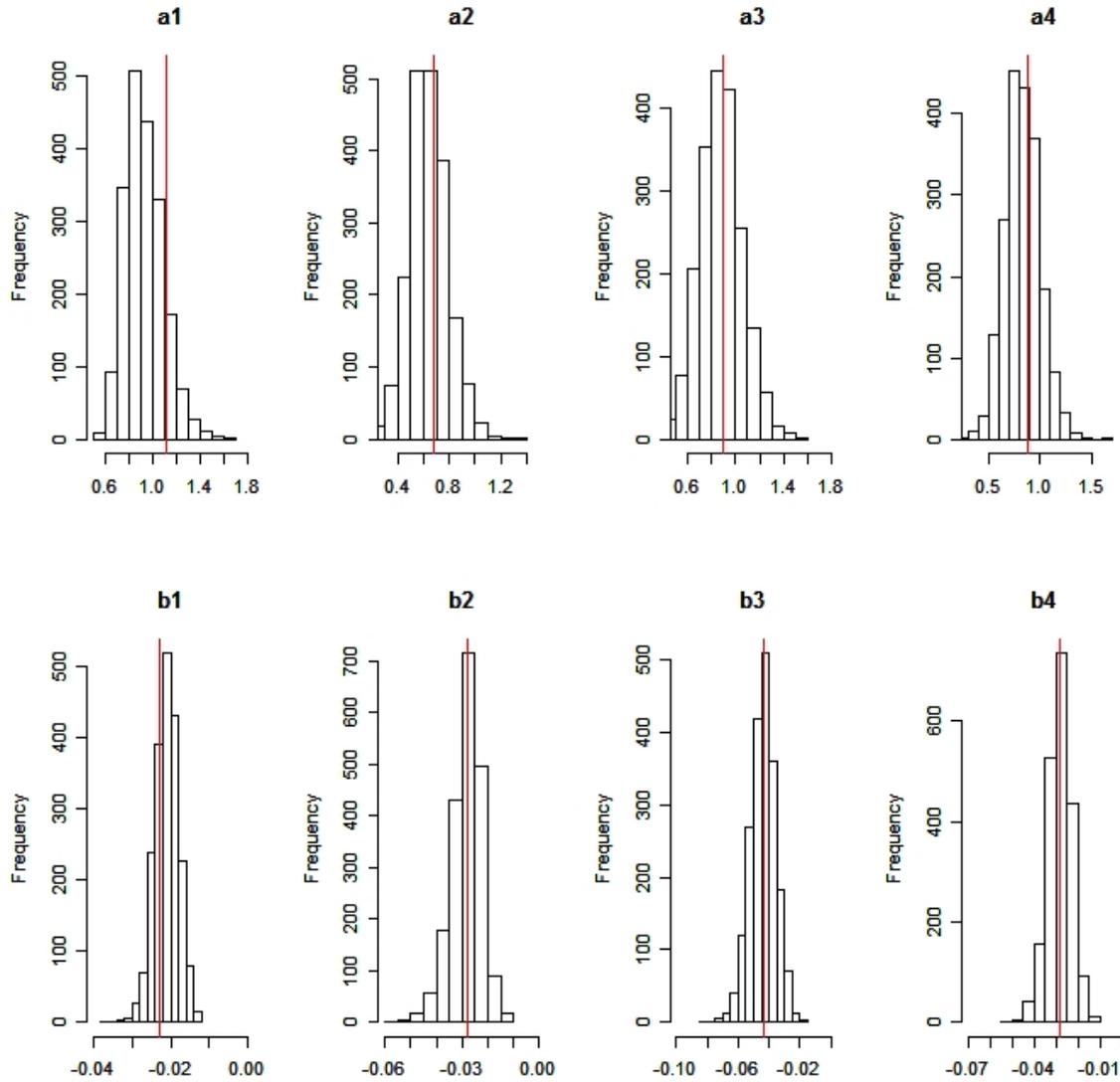
Por otro lado, con los datos de la Tabla 1<sup>2</sup> se realizaron las gráficas de log-verosimilitud perfil relativa con las cuales se obtuvieron los elementos necesarios para calcular los intervalos de verosimilitud para cada uno de los parámetros, además asegurar el haber obtenido el verdadero estimador máximo verosímil. Para esto, se creó una rejilla de 1000 valores alrededor de los verdaderos valores de los parámetros, de esta forma se obtuvo una rejilla lo suficientemente densa como para poder comparar y obtener los intervalos de verosimilitud del 15 %, que son los equivalentes a los intervalos del 95 % de confianza, con lo que se definen intervalos de aceptación y rechazo.

---

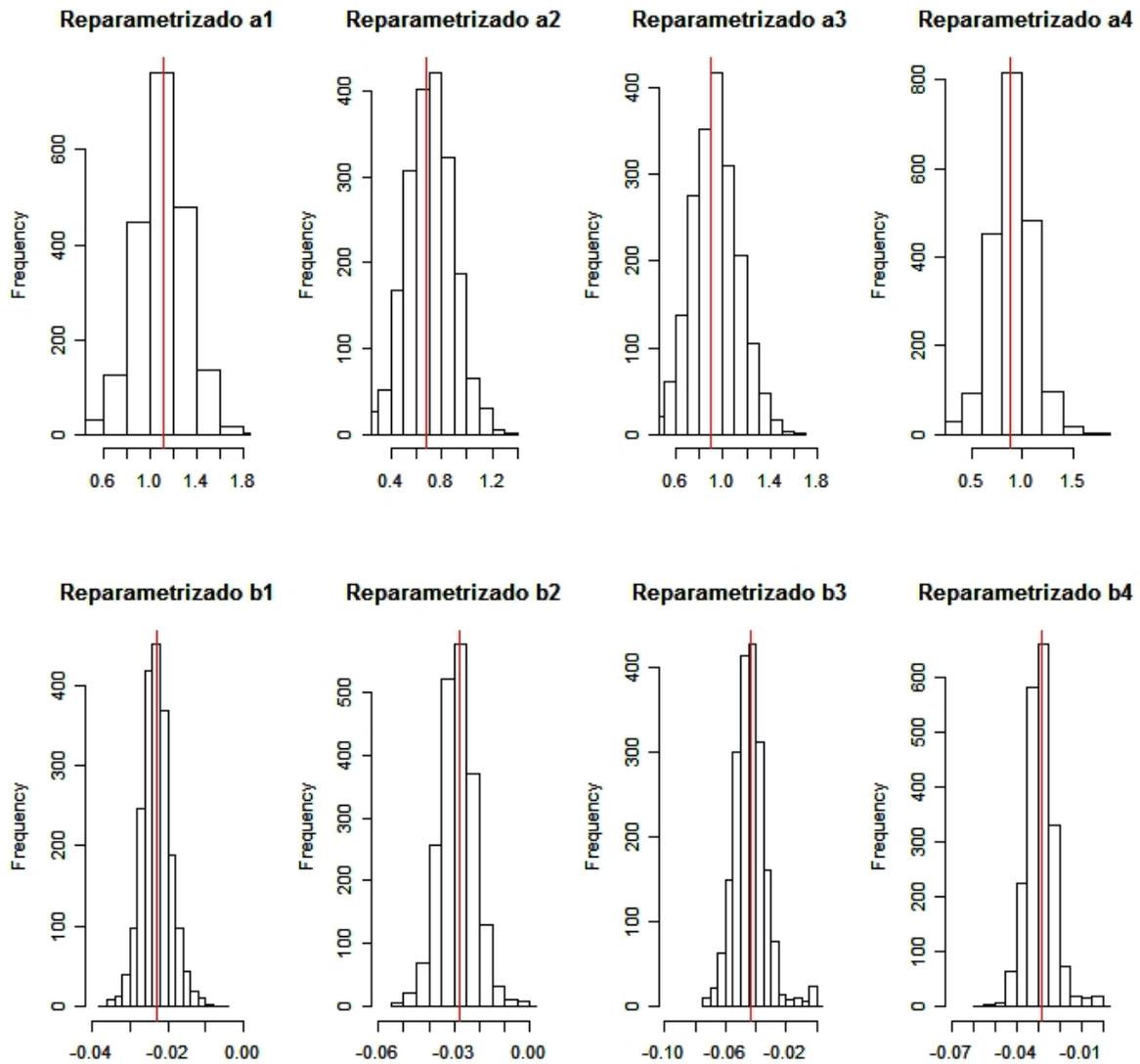
<sup>2</sup> Los datos contenidos en la Tabla 1, son resultado del estudio realizado a poblaciones de truchas toro en cuatro distintos afluentes de Estados Unidos (USA) durante el período de 1980 a 1993.

#### IV. RESULTADOS

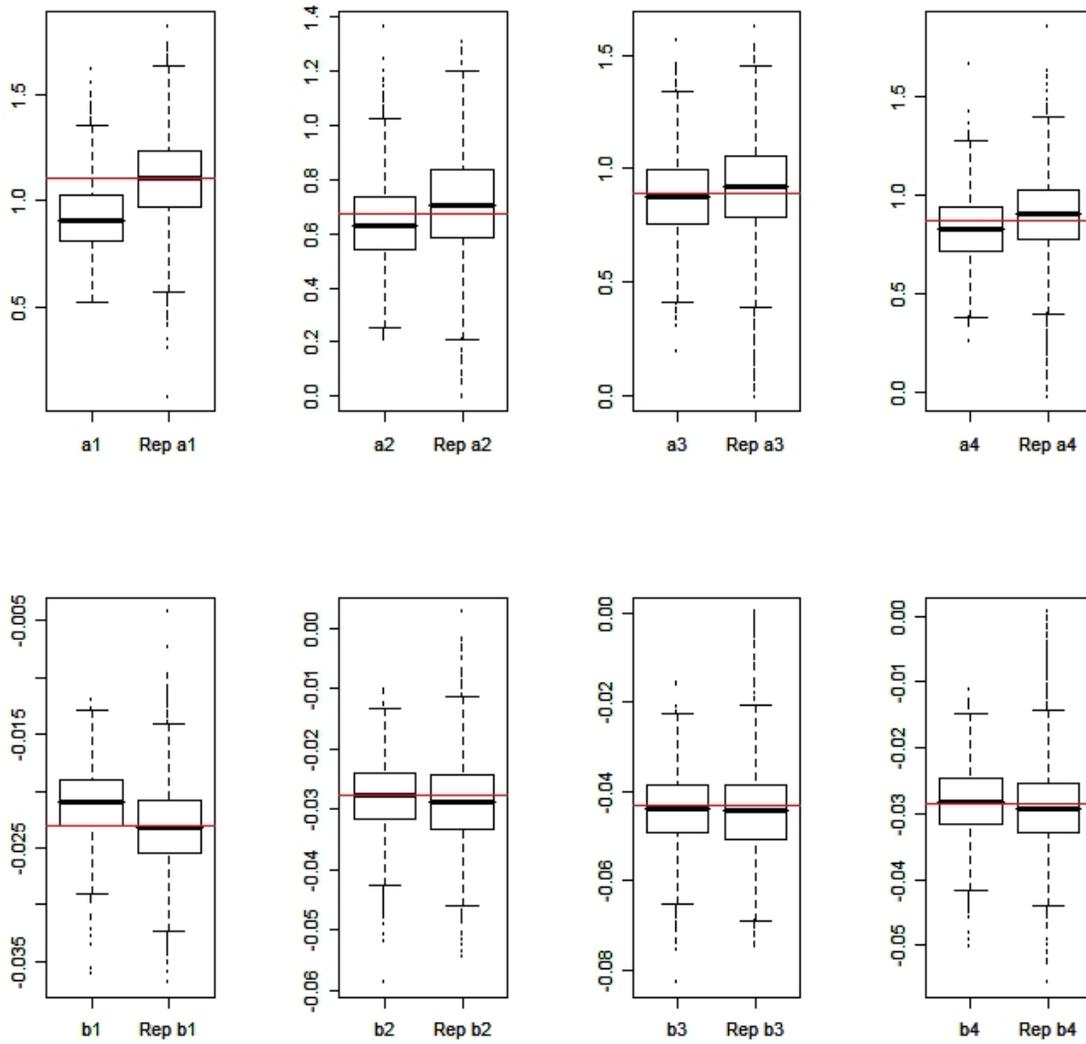
Gráfica 1.- Histogramas de frecuencia de los estimadores Bootstrap para los parámetros originales utilizados por Brian Dennis y Mark L. Taper. Las “a” son el parámetro de crecimiento exponencial estocástico y las “b” son el parámetro de crecimiento denso-dependiente.



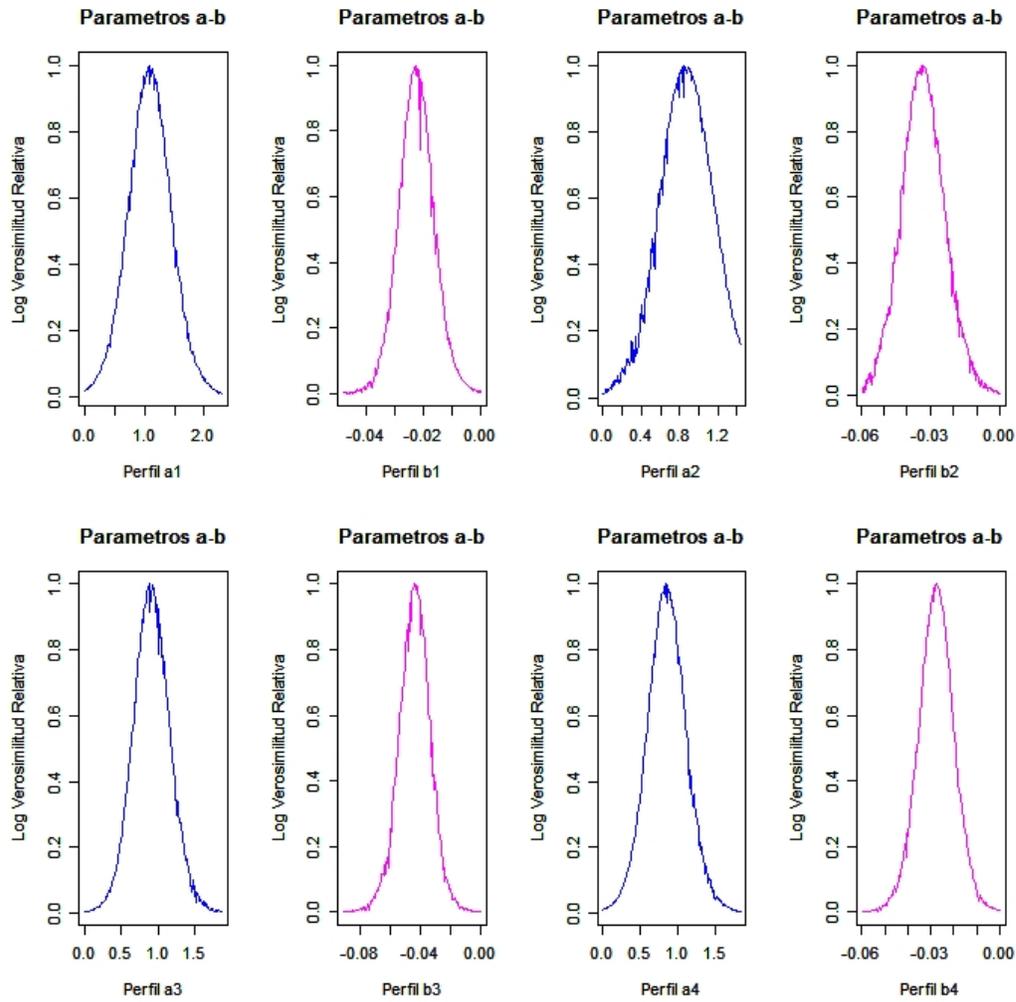
Gráfica 2.- Histogramas de frecuencia de los parámetros Bootstrap propuestos en el presente trabajo. Las “a” son los parámetros de crecimiento exponencial estocástico y las “b” son los parámetros de denso-dependencia con la transformación  $b = -\frac{a}{k}$ .



Gráfica 3.- Box-plot de cada parámetro y su contraparte, la línea roja indica el valor verdadero de cada parámetro.



Gráfica 4.-Gráficas de las funciones de Log Verosimilitud Perfil para los parámetros utilizados por Brian Dennis y Mark L. Taper.



Gráfica 5.-Gráficas de las funciones de Log Verosimilitud Perfil para los parámetros propuestos.

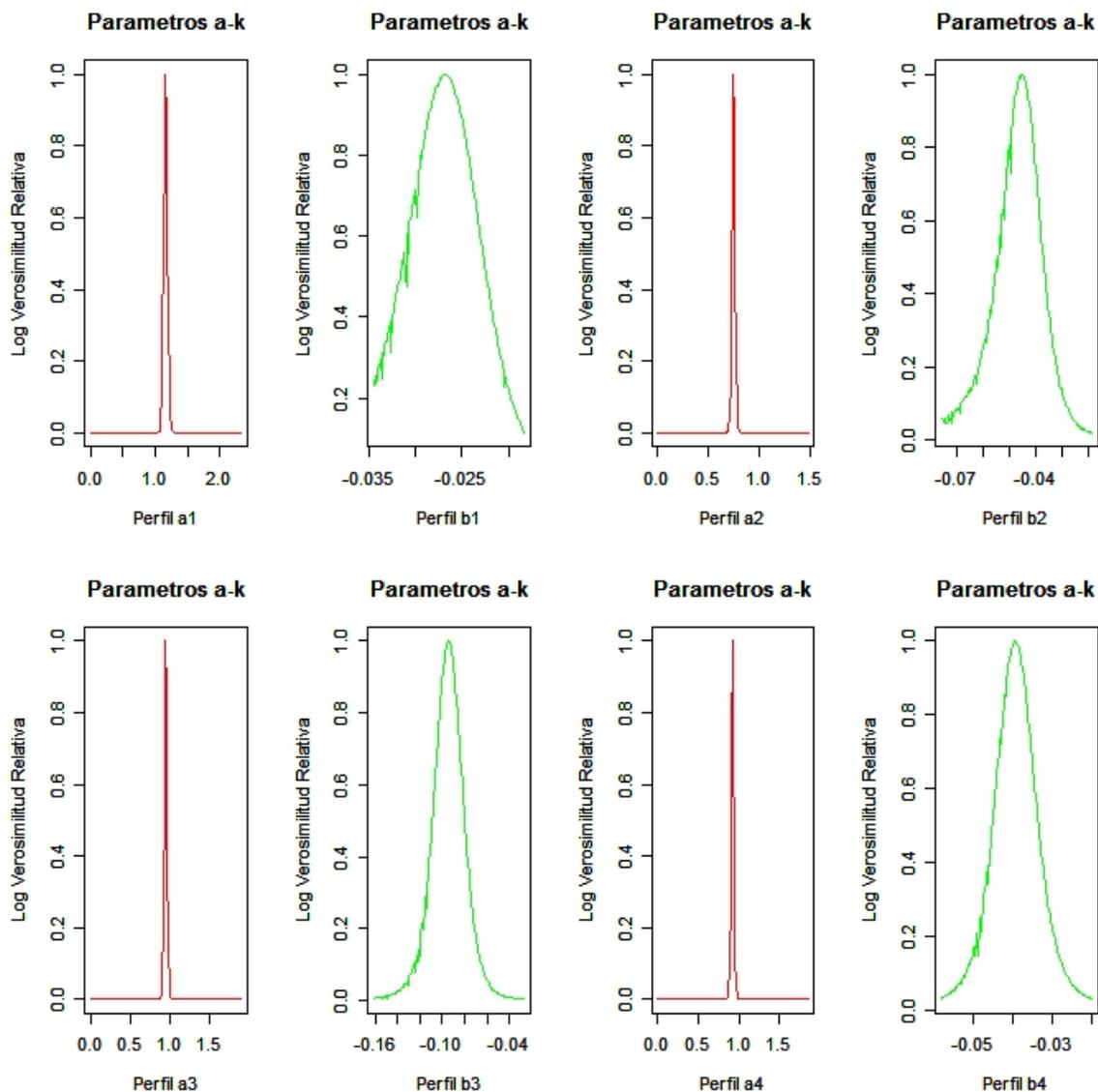


Tabla 2.- Tabla de resultados numéricos para la parametrización utilizada por Brian Dennis y Mark L. Taper.

	Verdad	Media	Mediana	Varianza	MSE	Sesgo en Media	Sesgo en Mediana
a1	1.11	1.1347542	1.1337162	0.0231941	0.0410741	0.1495772	0.1337162
a2	0.674	0.7228750	0.7183356	0.0198742	0.0992090	-0.1867821	-0.2816643
a3	0.893	0.9382005	0.932182	0.0243897	0.0289889	-0.0551869	-0.067817
a4	0.874	0.9151400	0.9076077	0.0241896	0.0327259	-0.0741676	-0.0923922
b1	-0.023	-0.0237084	-0.0236436	8.10E-06	1.0478543	0.0235452	-1.0236436
b2	-0.0276	-0.0298201	-0.0294997	3.30E-05	1.0599027	0.0284230	-1.0294997
b3	-0.043	-0.0453347	-0.0450509	5.30E-05	1.0921845	0.0449493	-1.0450509
b4	-0.0284	-0.0299277	-0.0297514	2.55E-05	1.0604136	0.0292499	-1.0297514

Tabla 3.- Tabla de resultados numéricos para la parametrización propuesta.

	Verdad	Media	Mediana	Varianza	MSE	Sesgo en Media	Sesgo en Mediana
Rep a1	1.11	1.1524531	1.1554378	0.0302171	0.0543780	0.1692230	0.1554378
Rep a2	0.674	0.7395269	0.734981	0.0285078	0.0987427	-0.1755588	-0.265018
Rep a3	0.893	0.9518888	0.9478821	0.0345854	0.0373017	-0.04296	-0.0521178
Rep a4	0.874	0.9283289	0.9263722	0.0316775	0.0370986	-0.062640	-0.0736277
Rep b1	-0.023	-0.0237534	-0.0237633	9.70E-06	1.0481010	0.023546	-1.0237633
Rep b2	-0.0276	-0.0300103	-0.0299277	4.27E-05	1.060793	0.0284282	-1.0299277
Rep b3	-0.043	-0.0455151	-0.0453354	6.95E-05	1.0927958	0.0449571	-1.0453354
Rep b4	-0.0284	-0.0300748	-0.0299499	3.16E-05	1.0608284	0.0292541	-1.0299499

## V. DISCUSIÓN DE RESULTADOS

Una vez obtenidos los estimadores por el método Bootstrap; los cuales tienden de manera asintótica a los estimadores de máxima verosimilitud; se analizó su distribución, por lo que era de esperarse que la media de la distribución de éstos estimadores fuera muy cercana a los parámetros de máxima verosimilitud reales, lo que se puede apreciar en los histogramas obtenidos para ambos conjuntos de parámetros.

La Gráfica 1 muestra los histogramas de frecuencia de los parámetros a-b según el modelo utilizado por Brian Dennis y Mark L. Taper, en éstos se aprecia como los estimadores Bootstrap forman una campana alrededor del estimador de máxima verosimilitud correspondiente; señalado por la línea vertical de color rojo. Así, de este modo la Gráfica 2 muestra los histogramas de frecuencia obtenidos con la re-parametrización a-k, de igual manera forman una campana alrededor del estimador de máxima verosimilitud. Observando ambos histogramas se concluye que las parametrizaciones correspondientes arrojan modelos que permiten realizar predicciones así como simulaciones para tamaños poblacionales muy eficaces, ya que en ambos casos se puede observar que la media poblacional de las estimaciones Bootstrap es muy cercana a los valores reales de máxima verosimilitud, que fueron obtenidos por Ricken en 1992<sup>[1]</sup>.

La Gráfica 3 muestra los Box-plot, que muestran el segundo cuartil (línea que aparece dentro del rectángulo), los denominados bigotes muestran el primer y tercer cuartil respectivamente, de este modo los puntos que aparecen fuera de estos bigotes son los que pueden ser tomados como datos atípicos, al obtener los Box-plot se colocó una línea horizontal de color rojo indicando donde se encuentra el verdadero valor del estimador de máxima verosimilitud, de este modo se pudo observar la distancia de éste con respecto a la mediana muestral (segundo cuartil), la primera caja de cada gráfico es el box-plot de la parametrización a-b y la segunda caja es la referente a la parametrización a-k, es claro observar que ambas parametrizaciones arrojan una cantidad muy similar de datos fuera de los bigotes y la mediana de ambas parametrizaciones está muy cerca del valor real.

Observando las Tablas 2 y 3 se aprecia que el sesgo de las medias difiere por décimas entre los parámetros de una parametrización y otra, éste fluctúa a favor de un modelo a otro por lo que el sesgo no muestra preferencia hacia alguna de las parametrizaciones. Por otro lado la varianza se inclina a favor de la parametrización a-b en orden de milésimas y centésimas, de estos dos valores se obtiene el error cuadrático medio en el cual no se observan diferencias significativas entre ambas parametrizaciones.

## VI. CONCLUSIONES

Al realizar las gráficas de la log verosimilitud de los modelos se observó que estas tienen una cresta muy “achatada” lo que afecta directamente al método de optimización (Nelder-Mead) utilizado para calcular el valor del estimador de máxima verosimilitud haciéndolo tener una fuerte dependencia al punto inicial y causando que arroje máximos locales privándonos de posible información más precisa, se utilizó el método Nelder-Mead de optimización ya que fue el utilizado por Brian Dennis y Mark L. Taper, además de ser el más popular en este tipo de estudios; se probó con otros dos métodos de optimización resultando mejor el método Nelder-Mead.

Al observar los gráficos y los datos contenidos en las tablas se concluye que no existe diferencia significativa entre una parametrización y otra; el uso del modelo propuesto por Ricken o el modelo propuesto en el presente trabajo depende del fin del estudio. Por ejemplo: si se desea estudiar la capacidad de carga del sistema es preferible utilizar la parametrización a-k (propuesta en el presente estudio) y esto evita realizar la transformación, si el fin del estudio es poder analizar los parámetros de densidad-dependencia se utilizará la parametrización a-b (de Ricken).

## II. REFERENCIAS BIBLIOGRÁFICAS

- 1.- Brian Dennis; Mark I. Taper(1994). Density Dependence in time series Observation of Natural Population: Estimation and Testing, Ecological Monograph
- 2.- <http://cran.r-project.org/doc/contrib/Owen-TheRGuide.pdf> (bibitem)
- 3.- Brian Dennis; Patricia L. Munholland; J. Michael Scott (1991). Estimation of Growth and Extinction Parameters for Endangered Species, Ecological Monograph.
- 4.- [http://cran.r-project.org/doc/contrib/rdebuts\\_es.pdf](http://cran.r-project.org/doc/contrib/rdebuts_es.pdf) (bibitem)
- 5.- Dra. Diana M. Kelmansky (2006). Introducción al lenguaje R: Dpto. de Matemáticas-Instituto de Cálculo. (bibitem)
- 6.- José Miguel Ponciano Castañeros (2004). Estimation of density dependence process noise and observation error: a comparison of modified maximum likelihood.
- 7.- Edgar acuña. Intervalo de Confianza Bootstrapping corregido por sesgo y acelerado (BCA). (bibitem)
- 8.- Brian Dennis; W. P. Kemp; M. L. Taper (1998). Joint Density Dependence. Ecology Vol. 79.

## Apéndice

### Tablas

Tabla 1.- Tabla de observaciones de tamaño poblacional por año de trucha toro en cuatro afluentes de Estados Unidos.

Año	Morrison	Granite	Lodgepole	Ole
1980	75	34	14	19
1981	32	14	18	19
1982	86	34	23	51
1983	67	31	23	35
1984	38	47	23	26
1985	99	24	20	30
1986	52	37	42	36
1987	49	34	21	45
1988	50	32	19	59
1989	63	31	43	21
1990	24	21	12	20
1991	45	20	9	23
1992	17	16	13	16
1993	14	9	9	19

### Funciones

Función de log-Verosimilitud para los parámetros a-b

```
verosimilitud<-function(params, vect){  
  d<-dim(vect); #Obtiene las dimensiones de la matriz dada.  
  m<-d[1]; #Obtiene el número de regiones analizadas para la verosimilitud.  
  q<-d[2]-1; #Obtiene el número de conteos.  
  s<-0; #Inicializa la suma de "s", que es uno de los dos sumandos  
  utilizados para obtener la log-verosimilitud.  
  a<-params[1:m]; # coloca los primeros m terminos en a  
  b<-params[(m+1):length(params)];# coloca los ultimos terminos en b
```

```

Y<-matrix(rep(0,(m*q)),nrow=m,ncol=q);#Crea la matriz Y para luego
    rellenarla con la diferencia de logaritmos de la matriz.

X<-log(vect); #genera la matriz de logaritmos de la matriz dada.

R<-matrix(rep(0,q*m),nrow=m,ncol=q); #Genera la matriz R que será usada para obtener la
matriz de varianzas y covarianzas.

for(i in 1:q){

    Y[,i] <- X[,i+1]-X[,i]; #Rellena Y con la resta consecutiva de los logaritmos de la matriz dada.

    for(j in 1:m){

        R[j,i]<- Y[j,i]-a[j]-b[j]*vect[j,i]; #Crea la matriz R que nos ayudará a generar la matriz de
varianzas y covarianzas.

    }

}

Sigma <- (R%*%t(R))/q; #Crea la matriz de varianzas y covarianzas según la verosimilitud del
modelo denso-dependiente.

Det.part <- -(q/2)*log(det(Sigma));

for(t in 1:q){

    s <- s + t(Y[,t]-a-diag(vect[,t])%*%b)%*%solve(Sigma)%*%(Y[,t]-a-diag(vect[,t])%*%b);

}

negloglike <- -Det.part +(1/2)*s; # Esta parte calcula el negativo de la log-verosimilitud.

return(negloglike);

}

```

Función de log-Verosimilitud para los parámetros a-k

```

verosimilitudk<-function(params, vect){

    d<-dim(vect); #Obtiene las dimensiones de la matriz dada.

```

```

m<-d[1]; #Obtiene el número de regiones analizadas para la verosimilitud.

q<-d[2]-1; #Obtiene el número de conteos.

s<-0; #Inicializa la suma de "s", que es uno de los dos sumandos utilizados para obtener la log-
verosimilitud.

a<-params[1:m]; # coloca los primeros m terminos en a

k<-params[(m+1):length(params)];# coloca los ultimos m terminos en k

Y<-matrix(rep(0,(m*q)),nrow=m,ncol=q);#Crea la matriz Y para luego rellenarla con la
diferencia de logaritmos de la matriz.

X<-log(vect); #genera X que es la matriz de logaritmos de la matriz dada.

R<-matrix(rep(0,q*m),nrow=m,ncol=q); #Genera la matriz R que será usada para obtener la
matriz de varianzas y covarianzas.

for(i in 1:q){

  Y[,i] <- X[,i+1]-X[,i]; #Rellena Y con la resta consecutiva de los logaritmos de la matriz dada.

  for(j in 1:m){

    R[j,i]<- Y[j,i]-a[j]-(-a[j]/k[j])*vect[j,i]; #Crea la matriz R que nos ayudará a generar la matriz
de varianzas y covarianzas.

  }

}

Sigma <- (R%*%t(R))/q; #Crea la matriz de varianzas y covarianzas según la verosimilitud del
modelo denso-dependiente.

Det.part <- -(q/2)*log(det(Sigma));

for(t in 1:q){

  s <- s + t(Y[,t]-a-diag(vect[,t])%*%(-a/k))%*%solve(Sigma)%*%(Y[,t]-a-diag(vect[,t])%*%(-
a/k));

}

negloglike <- -Det.part +(1/2)*s; # Esta parte calcula el negativo de la log-verosimilitud.

return(negloglike);

}

```

Función de verosimilitud Perfil para a-b

```
verosimilitud.perfil <- function(params, vect){  
  m<-vect[length(vect)-1]; #Obtiene el numero de regiones analizadas para la verosimilitud.  
  q<-vect[length(vect)]; #Obtiene el numero de conteos.  
  f<-c(vect[length(vect)-3],vect[length(vect)-2]);  
  vect<-matrix(vect[1:(m*(q+1))],ncol=(q+1));  
  s<-0; #Inicializa la suma de "s", que es uno de los dos sumandos utilizados para obtener la log-  
verosimilitud.  
  params[(f[1])] <- g[(f[2])];  
  a<-params[1:m]; # coloca los primeros m terminos en a  
  b<-params[(m+1):length(params)];# coloca los ultimos m terminos en b  
  Y<-matrix(rep(0,(m*q)),nrow=m,ncol=q);#Crea la matriz Y para luego rellenarla con la  
diferencia de logaritmos de la matriz.  
  X<-log(vect); #genera X que es la matriz de logaritmos de la matriz dada.  
  R<-matrix(rep(0,q*m),nrow=m,ncol=q); #Genera la matriz R que ser  usada para obtener la  
matriz de varianzas y covarianzas.  
  for(i in 1:q){  
    Y[,i] <- X[,i+1]-X[,i]; #Rellena Y con la resta consecutiva de los logaritmos de la matriz dada.  
    for(j in 1:m){  
      R[j,i]<- Y[j,i]-a[j]-b[j]*vect[j,i]; #Crea la matriz R que nos ayudar  a generar la matriz de  
varianzas y covarianzas.  
    }  
  }  
  Sigma <- (R%*%t(R))/q; #Crea la matriz de varianzas y covarianzas seg n la verosimilitud del  
modelo denso-dependiente.
```

```

Det.part <- -(q/2)*log(det(Sigma));

for(t in 1:q){

  s <- s + t(Y[,t]-a-diag(vect[,t])%*%b)%*%solve(Sigma)%*%(Y[,t]-a-diag(vect[,t])%*%b);

}

negloglike <- -Det.part +(1/2)*s; # Esta parte calcula el negativo de la log-verosimilitud.

return(negloglike);

}

```

Función de Verosimilitud Perfil para a-k

```

verosimilitudk.perfil <- function(params, vect){

  m <- vect[length(vect)-1]; #Obtiene el numero de regiones analizadas para la verosimilitud.

  q <- vect[length(vect)]; #Obtiene el numero de conteos.

  f <- c(vect[length(vect)-3],vect[length(vect)-2]);#nos dice que parámetros es el que se está
estudiando y en que

  #posición de la rejilla va.

  vect <- matrix(vect[1:(m*(q+1))],ncol=(q+1)); #Recomierte el vector en la matriz de datos.

  s <- 0; #Inicializa la suma de "s", que es uno de los dos sumandos utilizados para obtener la
log-verosimilitud.

  params[[f[1]]] <- gk[[f[2]]];

  a <- params[1:m]; # coloca los primeros m terminos en a

  k <- params[(m+1):length(params)];# coloca los ultimos m terminos en b

  Y <- matrix(rep(0,(m*q)),nrow=m,ncol=q);#Crea la matriz Y para luego rellenarla con la
diferencia de logaritmos de la matriz.

  X <- log(vect); #genera X que es la matriz de logaritmos de la matriz dada.

```

```
R <- matrix(rep(0,q*m),nrow=m,ncol=q); #Genera la matriz R que será usada para obtener la
matriz de varianzas y covarianzas.
```

```
for(i in 1:q){
```

```
  Y[,i] <- X[,i+1]-X[,i]; #Rellena Y con la resta consecutiva de los logaritmos de la matriz dada.
```

```
  for(j in 1:m){
```

```
    R[j,i]<- Y[j,i]-a[j]+(a[j]/k[j])*vect[j,i]; #Crea la matriz R que nos ayudará a generar la matriz
de varianzas y covarianzas.
```

```
  }
```

```
}
```

```
Sigma <- (R%*%t(R))/q; #Crea la matriz de varianzas y covarianzas según la verosimilitud del
modelo denso-dependiente.
```

```
Det.part <- -(q/2)*log(det(Sigma));
```

```
for(t in 1:q){
```

```
  s <- s + t(Y[,t]-a-diag(vect[,t])%*%(-a/k))%*%solve(Sigma)%*%(Y[,t]-a-diag(vect[,t])%*%(-
a/k));
```

```
}
```

```
negloglike <- -Det.part +(1/2)*s; # Esta parte calcula el negativo de la log-verosimilitud.
```

```
return(negloglike);
```

```
}
```