UNIVERSIDADE DE SÃO PAULO

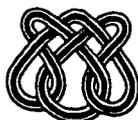## Instituto de Ciências Matemáticas e de Computação

**Inference for the Richards Growth Model
Using Box and Cox Transformation
and Bootstrap Techniques**

Selene Loibel
João B. R. do Val
Marinho G. Andrade

Nº 75

NOTAS

Série Estatística

São Carlos – SP

Ago./2002

Inference for the Richards Growth Model using Box and Cox Transformation and Bootstrap Techniques

Selene Loibel, João Bosco R. do Val
DT / FEEC / UNICAMP
P.B. 6101, 13081–970 - Campians /SP

Marinho G. Andrade
SCE / ICMC / USP
P.B. 668, 13560-970 - São Carlos / SP

19 de agosto de 2003

Resumo

Neste artigo consideramos o problema de identificação do modelo de crescimento populacional de Richards. Entre os modelos de crescimento populacional o modelo de Richards é considerado por ser suficientemente geral e incorporar como casos particulares os modelos logístico e Gompertz. A utilização deste modelo apresenta problema de identificabilidade principalmente em problemas com poucos dados o que é comum em modelamento de populações animais. Neste trabalho mostramos que o modelo de Richards pode ser interpretado como um modelo logístico quando aplicamos aos dados a transformação de potência de Box e Cox. Devido às particularidades do modelo de Richards, como a não linearidade, foi necessário considerar a função de verossimilhança profile no cálculo das estimativas clássicas pontuais. O uso da teoria assintótica para o cálculo de intervalos de confiança para os parâmetros é desaconselhável e a alternativa proposta para o cálculo de intervalos de confiança é a técnica boostrap paramétrico. A abordagem proposta foi testada com dados gerados e os resultados obtidos permitiram a comparação entre os intervalos de confiança estimados com a teoria assintótica e com a técnica bootstrap.

Palavras-chave: Modelo de Richards, Modelos de crescimento populacional, Função de verossimilhança profile, Transformação de Box e Cox. Bootstrap paramétrico.

Inference for the Richards Growth Model using Box and Cox Transformation and Bootstrap Techniques

Selene Loibel[a] João B. R. do Val[a,1] Marinho G. Andrade[b]

[a] *UNICAMP – Fac. de Engenharia Elétrica e de Computação, Depto. de Telemática, C.P. 6101, 13081-970, Campinas, SP, Brazil*

[b] *USP – ICMC, Depto. de Ciências de Computação e Estatística, C.P. 668, 13560-970, São Carlos, SP, Brazil*

Abstract

The paper considers the parameter identification of the Richards population growth model, a class of models that generalizes the logistic and the Gompertz growth model classes. The use of Richards class presents an identifiability problem that is worsened when handling small data samples, a situation often encountered in animal population modeling. This paper tackles the identifiability problem by applying the Box and Cox transformation to the data, leading to a model that can be interpreted as a logistic growth model. The approach is complemented with the profile maximum likelihood estimates of an intrinsic parameter of the Richards model, combined with the bootstrap technique to assure better interval estimation for the model parameters. Some tests with generated data are presented to illustrate the technique.

Keywords Richards growth model, profile likelihood function, Box and Cox transformations, parametric bootstrap.

1 Introduction

There are many reasons to employ modeling techniques when studying the growth of an animal population. Firstly, modeling is an important tool to understand how environmental uncertainties affect the population growth; secondly, these models can be used to forecast the population behavior as well as to estimate their extinction risk and other statistics connected with the

---

[1] Corresponding author. email: jbosco@dt.fee.unicamp.br Fax: +55(19)3289-1395 Phone: +55(19)3788-3762

population extinction (see Boyce (1992), Hertzler et al. (1997), Homes (2001) and Ludwig (1996)). For actual calculations, accurate parameter identification is necessary.

To identify the parameters of a given model two concepts should be considered: the parameters point estimate and the estimate interval. For parameter estimate, simple values (or estimates) are obtained for the set of data being observed, from the local properties of the model in the parameters space. These estimates can only be determined if the parameters can be identified in an unique way for a certain set of data, thus, being a necessary condition for that classical analysis. If the model parameters are not identifiable in an unique way, the model structure must be simplified until the identifiability is achieved.

An alternative approach is to estimate the model parameters by intervals. This procedure consists of estimating a probable region for the parameters and therefore it does not require a unique solution for the parameter estimates. When there is a good parameter identifiability, the interval estimate results in a narrow region (or in a narrow confidence interval) for the parameters. When the identifiability is poor, and there are one or more non-identifiable parameters, there will be a more extensive region for the parameter interval estimates with poor identifiability characteristics and for the parameters correlated to them.

In ecological models, poor identifiability of the parameters is common because in general the data used to estimate these parameters are scarce in relation to the complexity of the models (Homes, 2001). On those situations, estimates by intervals are more informative than simple point estimates. One of the approaches used to obtain interval estimates is the Bayesian inference (Gelman et al., 1995); that approach is interesting and would be applicable to very complex structural models (see Omlin and Reichert (1999), Reichert and Omlin (1997)). However, due to the complexity of natural systems, it is difficult to express the prior knowledge of theirs mechanisms. These facts increase the interest on techniques of interval estimates applied to ecological modeling. In this connection, the parametric bootstrap technique can be useful because small samples can be considered and it does not require model simplifications.

The Richards growth model (Richards, 1959) is a generalization of the logistic and Gompertz growth models. It was Initially applied in animal growth studies (weight versus age) and afterwards Fitzhugh (1974) suggested that it could be used in animal population growth modeling. the two special cases of Richards model are used in most of the works found in literature, due mainly, to the difficulties of making inferences from Richards non-linear model (see Ludwig (1996), Ludwig (1999)).

2

The difficulties of the asymptotic theory and the classical approach for the parameter identification of Richards growth model are related to the bad conditioning of the Fisher information matrix associated to the model. In order to circumvent these problems, this paper shows that Richards model can be interpreted as a logistic model when we apply the Box and Cox transformation to the data (Box and Cox, 1964). That leads to the method of profile maximum likelihood (Barndorff-Nielsen and Cox, 1994), which makes the inference of the parameters feasible. Due to the bad conditioning of Fisher information matrix and the relation with bad parameters estimates, the mere application of the profile maximum likelihood leads to imprecise results for the confidence intervals of model parameters. Sometimes they do not include the real values of the parameters due to the complexities of Richards model and also due to limited use of the asymptotic theory for small sample sizes. In order to overcome this problem the parametric bootstrap technique is considered here (see Efron and Tibshirani (1993), Wehrens et al. (2000)). Using two data sets, we show that the proposed method estimates precise confidence intervals for the model parameters.

The paper is organized as follows. In section 2 we present the Richards growth model as a model that generalizes the logistic and the Gompertz models. In section 3 we apply the Box and Cox transformation to obtain linear logistic model if a certain intrinsic parameter of the Richards model is known. In section 4 we develop the likelihood function for the model and show how the profile likelihood function can be used to obtain the whole set of parameter estimates. In section 5 we discuss the confidence intervals, in preparation for section 6 where the parametric bootstrap techniques are developed in some details. Finally, section 7 presents and discuss some numerical experiments and section 8 presents the conclusions of the paper.

2 The Richards growth model

Let us consider the discrete-time population exponential growth equation in a general form

$$N_{t+1} = N_t \exp\{r(N_t) + \epsilon_t\} \tag{1}$$

In particular, the Richards growth model can be written as

$$N_{t+1} = N_t \exp\left\{\frac{\rho}{q}\left[1 - \left(\frac{N_t}{K}\right)^q\right]\right\} \exp\{\epsilon_t\} \tag{2}$$

where ρ, q and K are some parameters to be detailed below. $\{\epsilon_t, t = 0, 1 \ldots\}$ is a stochastic process used to represent the variability of environmental factors. It is assumed that $\{\epsilon_t, t \geq 0\}$ has a Gaussian distribution with $E(\epsilon_t) = 0$,

$E(\epsilon_t^2) = \tau^{-1}$, $(\tau = 1/\sigma^2 > 0)$ and $E(\epsilon_t, \epsilon_{t+k}) = 0$ for all $k \neq 0$, namely, it is a white noise process.

The Richards growth rate is

$$r(N_t) = \frac{\rho}{q}\left[1 - \left(\frac{N_t}{K}\right)^q\right] \qquad (3)$$

The parameter ρ is the population intrinsic growth rate, that is, it is the difference between the birth and death rates of that population in question: $\rho = d - m$. The K parameter is the carrying capacity, or the limiting size for the population above which the growth rate becomes decreasing. This limiting size is imposed by environmental factors, basically food and space limitations. The parameter q represents a population intrinsic factor, that describes how the rate $r(N_t)$ decreases as N_t increases. That is, if $q > 1$, $r(N_t)$ has a slow decrease, if $q < 1$, $r(N_t)$ has a faster decrease and if $q = 1$ the rate $r(N_t)$ varies linearly with N_t.

The Richards growth rate is quite general because when varying the value of q we span a family of well known models. To later use, we can write the Richards rate equivalently as

$$r(N_t) = \alpha_1 + \alpha_2 N_t^q \qquad (4)$$

with $\alpha_1 = \rho/q$ and $\alpha_2 = -\rho/qK^q$, hence,

$$K = \left(\frac{-\alpha_1}{\alpha_2}\right)^{1/q} \text{ and } \rho = q\alpha_1 \qquad (5)$$

Remark 1 (Logistic and Gompertz growth models) *Richards growth model has two models as special cases that are used on many population studies: The logistic growth model and the Gompertz growth model.*

Logistic growth model: *With $q = 1$ in (3) we can write the following logistic growth rate*

$$r(N_t) = \rho\left[1 - \frac{N_t}{K}\right]$$

The logistic rate is written in a linear form as in (4) and (5) with q=1.

Gompertz growth model: *Gompertz growth rate is obtained by considering the Richards growth rate (3) in the limiting situation $q \to 0$, as follows*

$$\lim_{q\to 0} r(N_t) = \lim_{q\to 0} \frac{\rho}{q}\left[1 - \left(\frac{N_t}{K}\right)^q\right] = \lim_{q\to 0} \frac{\rho}{q}\left[1 - \exp\left\{q\ln\left(\frac{N_t}{K}\right)\right\}\right]$$

4

By applying the L'Hôpital rule, we have Gompertz growth rate as

$$\lim_{q \to 0} r(N_t) = \lim_{q \to 0}\left[-\rho \ln\left(\frac{N_t}{K}\right)\exp\left\{q\ln\left(\frac{N_t}{K}\right)\right\}\right] = -\rho\ln\left(\frac{N_t}{K}\right)$$

Analogously, the Gompertz growth rate is written in a linear form as

$$r(N_t) = \alpha_1 + \alpha_2\ln(N_t) \qquad (6)$$

with $\alpha_1 = \rho\ln(K)$ and $\alpha_2 = -\rho$; thus, $K = \exp\left\{-\alpha_0/\alpha_1\right\}$.

3 Inference of Richards growth model

The calculation of classical estimates for the Richards growth model parameters presents many difficulties. A problem is that the use of numerical methods (Newton, quasi-Newton etc.) for the likelihood function maximization is not possible since the Hessian matrix for this model may be singular. We show that this difficulty can be circumvented by the use of the Box and Cox transformation applied to the data and the calculation of the profile likelihood estimates (Barndorff-Nielsen and Cox, 1994) for the model. We set the shape parameter q to coincide with the Box and Cox power transformation parameter and by this the Richards growth model is interpreted as a logistic model. To show this, let us consider the Box and Cox power transformation given as

$$N_t^{(q)} = \begin{cases} (N_t)^q, & \text{if } q > 0 \\ \ln(N_t), & \text{if } q = 0 \end{cases} \qquad (7)$$

It is convenient to adopt the notation $N_t^{(q)} = (N_t)^q$, $q > 0$, and the logistic growth model associated to the transformed data is written as

$$N_{t+1}^{(q)} = N_t^{(q)}\exp\left\{\rho\left(1 - \frac{N_t^{(q)}}{K_{\mathrm{L}}}\right)\right\}\exp\{\epsilon_t^L\} \qquad (8)$$

By applying the power $1/q$ to (8), we get the Richards growth model as in (1)

$$N_{t+1} = N_t\exp\left\{\frac{\rho}{q}\left[1 - \left(\frac{N_t}{K_R}\right)^q\right]\right\}\exp\{\epsilon_t\}, q > 0 \qquad (9)$$

where we consider $K_R = K_L^{1/q}$ and $\epsilon_t = \epsilon_t^L/q$.

5

Denoting $Y_t^{(q)} = \ln(N_{t+1}^{(q)}/N_t^{(q)})$ the equation (8) can be written in linear form as

$$Y_t^{(q)} = \alpha_1 + \alpha_2 N_t^{(q)} \epsilon_t^L \tag{10}$$

where $\alpha_1 = \rho$, $\alpha_2 = -\rho/K_L$ and ϵ_t represents a noise with Gaussian distribution with $E(\epsilon_t^L) = 0$, $E((\epsilon_t^L)^2) = \tau_0^{-1}$ and $E(\epsilon_t^L \epsilon_{t+k}^L) = 0$, for all $k \neq 0$.

4 The profile likelihood function (PLF)

We consider the logistic growth model equation (10) to estimate the Richards parameters using the Box and Cox transformation, combining with the PLF method. When writing the model in the form (10), there is no numerical difficulties to estimate α_0 and α_1 for a fixed value of q. The next step is to find an estimate $\widehat{q}$ for q that maximizes the PLF. Having the corresponding parameters estimates $\widehat{\alpha}_1$ and $\widehat{\alpha}_2$, we can calculate the parameter estimates $\widehat{\rho}$ and $\widehat{K}$ by using the basic relations

$$\widehat{\rho} = \widehat{\alpha}_1 \text{ and } \widehat{K} = \left(-\frac{\widehat{\alpha}_1}{\widehat{\alpha}_2}\right)^{1/\widehat{q}} \tag{11}$$

In the previous section we consider $\epsilon_t^L \sim \text{Normal}(0, \tau_0^{-1})$, therefore, to obtain the PLF we use the fact that the probability density function of $Y_t^{(q)}|N_t^{(q)}$ is Gaussian with mean $\alpha_1 + \alpha_2 N_t^{(q)}$ and variance $\text{var}(z_t) = \tau_0^{-1}$. Hence,

$$P(Y_t^{(q)}|N_t^{(q)}) = \left(\frac{\tau_0}{2\pi}\right)^{1/2} \exp\left\{-\frac{\tau_0}{2}(y_t^{(q)} - \alpha_1 - \alpha_2 N_t^{(q)})^2\right\}$$

The aim is to estimate the values that maximize the original data likelihood function, and we need to refer to the original variables considering the Jacobian of the transformation

$$P(N_{t+1}|N_t) = P(N_{t+1}^{(q)}|N_t^{(q)})\left|\frac{dN_{t+1}^{(q)}}{dN_{t+1}}\right| = P(N_{t+1}^{(q)}|N_t^{(q)})q N_{t+1}^{q-1} \tag{12}$$

But according to (10) we have that

$$P(N_{t+1}^{(q)}|N_t^{(q)}) = P(Y_t^{(q)}|N_t^{(q)})\left|\frac{dY_t^{(q)}}{dN_{t+1}^{(q)}}\right| \tag{13}$$

By substituting (4) in (13) and using the relation $Y_t^{(q)} = \ln(N_{t+1}^{(q)}/N_t^{(q)})$ to

calculate the Jacobian in (13) we have the density $P(N_{t+1}^{(q)}|N_t^{(q)})$ given as

$$P(N_{t+1}^{(q)}|N_t^{(q)}) = \left(\frac{\tau_0}{2\pi}\right)^{1/2} \exp\left\{-\frac{\tau_0}{2}(y_t^{(q)} - \alpha_1 - \alpha_2 N_t^{(q)})^2\right\}\frac{1}{N_{t+1}^q} \qquad (14)$$

Finally, we can write the density $P(N_{t+1}|N_t)$ as

$$P(N_{t+1}|N_t) = \left(\frac{\tau_0}{2\pi}\right)^{1/2} \exp\left\{-\frac{\tau_0}{2}(y_t^{(q)} - \alpha_1 - \alpha_2 N_t^{(q)})^2\right\}\frac{q}{N_{t+1}} \qquad (15)$$

Denoting the data vector by $\boldsymbol{D} = (N_1, \ldots, N_n)$, the transformed data vectors by $\boldsymbol{Y}^{(q)} = (Y_1^{(q)}, \ldots, Y_n^{(q)})$ and $\boldsymbol{N}^{(q)} = (N_1^{(q)}, \ldots, N_n^{(q)})$ with $q > 0$ and the parameters vector by $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$, the likelihood function is given as

$$L(q, \boldsymbol{\alpha}, \tau|\boldsymbol{D}) = \left(\frac{\tau_0}{2\pi}\right)^{(n-1)/2} \exp\left\{-\frac{\tau_0}{2}(\boldsymbol{Y}^{(q)} - \boldsymbol{X}^{(q)}\boldsymbol{\alpha})'(\boldsymbol{Y}^{(q)} - \boldsymbol{X}^{(q)}\boldsymbol{\alpha})\right\}\frac{q^{n-1}}{\prod_{t=1}^{n-1} N_{t+1}}$$
$$(16)$$

where

$$\boldsymbol{X}^{(q)} = \begin{pmatrix} 1 & N_1^{(q)} \\ \vdots & \vdots \\ 1 & N_n^{(q)} \end{pmatrix} = \begin{bmatrix} 1 & \boldsymbol{N}^{(q)} \end{bmatrix}$$

The values of $\boldsymbol{\alpha}$ and τ that maximize the function in (16) also maximize the logarithm of that function. Denoting $l(q, \boldsymbol{\alpha}, \tau_0|\boldsymbol{D}) = \ln[L(q, \boldsymbol{\alpha}, \tau_0|\boldsymbol{D})]$, we have that

$$l(q, \boldsymbol{\alpha}, \tau_0|\boldsymbol{D}) = \frac{(n-1)}{2}[\ln(\tau_0) - \ln(2\pi)] - \frac{\tau_0}{2}(\boldsymbol{Y}^{(q)} - \boldsymbol{X}^{(q)}\boldsymbol{\alpha})'(\boldsymbol{Y}^{(q)} - \boldsymbol{X}^{(q)}\boldsymbol{\alpha}) +$$
$$+ (n-1)\ln(q) - \sum_{t=1}^{n-1} \ln(N_{t+1})$$
$$(17)$$

To calculate the parameters estimates of $\boldsymbol{\alpha}$, τ_0 e q we consider (17) as the PLF, which we denoted by $l_p(q|\boldsymbol{D}) = l(q, \widehat{\boldsymbol{\alpha}}_q, \widehat{\tau}_{0q}|\boldsymbol{D})$, where $\widehat{\boldsymbol{\alpha}}_q$ e $\widehat{\tau}_{0q}$ are the likelihood maximum estimates for $\boldsymbol{\alpha}$, τ_0 for a fixed q. By calculating $l_p(q|D)$ for q in an interval, we could identify the maximum of the function and the likelihood maximum estimate $\widehat{q}$ for q. Besides, we could calculate a confidence interval for q according to the relation

$$2[l_p(\widehat{q}) - l_p(q)] \to \chi_{\nu_q}^2 \qquad (18)$$

where ν_q is the number of independent components in q. More details on the method can be found in Barndorff-Nielsen and Cox (1994).

7

5 Estimating the confidence intervals for ρ, K and τ

Let us consider the relation between the parameter vector $\gamma = (\alpha_1, \alpha_2, \tau_0, q_0)$ and the parameter vector of interest $\theta = (\rho, K, \tau, q)$. The relations are: $\alpha_1 = \rho/q$, $\alpha_2 = -\rho/qK^q$, $\tau_0 = \tau/q^2$, $q_0 = q$. We can estimate the confidence intervals for $\theta_q = (\rho_q, K_q, \tau_q)$ by using the asymptotic normality property of the profile maximum likelihood estimators namely $\widehat{\theta}_q \sim N\left(\theta_q, I_q(\widehat{\theta}_q)^{-1}\right)$ where $I_q(\widehat{\theta}_q)$ is the observed Fisher information matrix for θ given q.

Denoting by $I(\widehat{\gamma})$ the observed Fisher information matrix for the parameter vector γ, and by $I(\widehat{\theta})$ the observed Fisher information matrix for θ, we can calculate $I(\widehat{\theta})$ using the Jacobian given by the matrix $J_{\gamma\theta}$ and the following relations

$$I(\widehat{\theta}) = J'_{\gamma\theta}\, I(\widehat{\gamma}) J_{\gamma\theta} = \begin{pmatrix} I_{\widehat{\theta}_q \widehat{\theta}_q} & I_{\widehat{\theta}_q q} \\ I_{q\widehat{\theta}_q} & I_{qq} \end{pmatrix} \tag{19}$$

$$J_{\gamma\theta} = \frac{\partial \gamma}{\partial \theta} = \begin{pmatrix} 1/q & 0 & 0 & -\rho/q^2 \\ -1/qK^q & \rho/K^{q+1} & 0 & \rho/(q^2K^q) + \rho\ln(K)/(qK^q) \\ 0 & 0 & 1/q^2 & -2\tau/q^3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{20}$$

Due to the bad conditioning of the Fisher information matrix it is not possible to calculate $I(\widehat{\theta})^{-1}$ directly, but we can use the relation $\widehat{V}(\widehat{\theta}_q) = [I_{\widehat{\theta}_q\widehat{\theta}_q} - I_{\widehat{\theta}_q q}I_{qq}^{-1}I'_{\widehat{\theta}_q q}]^{-1}$ to calculate the approximate variance of θ_q (see (Seber and Wild, 1989)).

Remark 2 (The difficulties of the classical inference) *One of the difficulties presented in the methods just described is that the samples available for the study of animal population growth are generally small. That makes impossible the use of the asymptotic theory in the sense that we could not consider the asymptotic normality property of both the maximum likelihood estimators and the profile maximum likelihood estimators. This fact makes the inference poor since we cannot have good estimates for the parameters standard deviation but only the point estimates for the parameters. With a small data sample to deal with, one possible option is the use of the bootstrap technique. Briefly, this technique is a re-sampling procedure for estimate precision measurements such as standard deviation, bias and confidence intervals. The advantage of this approach is to provide each of the parameter estimates: mean, median, correlation coefficient, etc. We will apply the bootstrap technique to the model*

that we have developed so far, when a small data sample is available.

6 The parametric bootstrap

According to bootstrap pattern literature, we denoted by X a random variable with cumulative distribution function $F(X, \theta)$. We also denoted the maximum likelihood estimate of θ by $\hat{\theta} = g(X)$.

The estimated confidence interval of θ is usually obtained from the estimator $\hat{\theta}$ by considering the probability distribution of $\hat{\theta} - \theta$. If s_α denotes the α-percentile of the distribution, then the confidence interval for θ is based on the probability statement $P(s_{\alpha/2} \leq \hat{\theta} - \theta \leq s_{1-\alpha/2}) = 1 - \alpha$ which, after rewriting, leads to the interval $\hat{\theta} - s_{1-\alpha/2} \leq \theta \leq \hat{\theta} - s_{\alpha/2}$. Note that to calculate the interval it is necessary to know or, at least to be able to approximate the distribution of $\hat{\theta} - \theta$. The hypothesis of asymptotic normality assumes that $\hat{\theta} \sim N(\theta, I^{-1}(\theta))$ where $I(\theta)$ is the Fisher information matrix (see section 5), or alternatively considers the approximation $(\hat{\theta} - \theta) \sim t_{n-p}$ where t_{n-p} represents the distribution t-student with $n - p$ degrees of freedom (p is the total number of unknown parameters to be estimated).

The bootstrap method is based on the idea that the variability of $\hat{\theta}$ around θ is similar to the variability of the bootstrap estimator $\hat{\theta}^*$ around $\hat{\theta}$. In some cases the bootstrap estimator can be calculated analytically, but in general re-samplings are necessary. In these cases, the procedure known as parametric bootstrap consists of generating samples from $F(X, \hat{\theta})$. The B samples generated are denoted by $X^*(b), b = 1 \ldots, B$ and the bootstrap estimates of θ for each one of those samples are denoted by $\theta^*(b), b = 1, \ldots, B$. With the boostrap estimates, we can calculate the standard error of $\hat{\theta}$ given by

$$\widehat{se}^*(B) = \left\{ \frac{\sum_{b=1}^{B} (\theta^*(b) - \hat{\theta}^*)^2}{B - 1} \right\}^{1/2} \tag{21}$$

where $\hat{\theta}^*$ is given by

$$\hat{\theta}^* = \frac{\sum_{b=1}^{B} \theta^*(b)}{B} \tag{22}$$

The ideal bootstrap estimate of the standard error would be obtained when $B \rightarrow \infty$, but the amount of computing time to evaluate the bootstrap estimates increases linearly with B. It is obvious that the precision of the standard error estimate must be as good as any other quantity of interest and therefore we expect that the estimate presents a small bias and small standard deviation. The ideal bootstrap estimate $\widehat{se}^*(\infty)$ has the smallest possible deviation among unbiased estimates.

One manner of choosing a satisfactory value for B is to calculate the coefficient of variation of $\widehat{se}^*(B)$. The increased variability due to stopping after B boostrap replications is reflected in an increased coefficient of variation. According to Efron and Tibshirani (1993), more than 200 replications are needed for estimating the standard error and much larger values of B are required for bootstrap confidence intervals.

6.1 Bootstrap confidence intervals

The bootstrap technique allows precise intervals without the asymptotic normality assumption, given as $(\widehat{\theta}-\theta) \sim N(0, I^{-1})$ or the approximation $(\widehat{\theta}-\theta) \sim t_{n-p}$. Many methods have been proposed to estimate the bootstrap confidence interval; there is a review in Efron and Tibshirani (1993) and a more recent review can be found in Wehrens et al. (2000). Among the proposed techniques, the most precise ones are: 1) Percentile method 2) Corrected percentile method BC_α (bias-corrected and accelerated) and 3) ABC (approximate bootstrap confidence intervals).

Percentile method: The confidence interval based on the bootstrap distribution, comes from the assumption that the distribution of $(\widehat{\theta} - \theta)$ is approximated by the $(\widehat{\theta}^* - \theta)$ distribution. Consequently, the percentiles of the distribution of $(\widehat{\theta} - \theta)$ are also approximated by the bootstrap distribution percentiles of $(\widehat{\theta}^* - \theta)$. In particular, the percentiles $s_{\alpha/2}$ and $s_{1-\alpha/2}$ of the $(\widehat{\theta} - \theta)$ distribution that are used for the classical confidence interval, namely

$$\widehat{\theta} - s_{1-\alpha/2} \leq \theta \leq \widehat{\theta} - s_{\alpha/2} \tag{23}$$

are substituted by the bootstrap approximation percentiles $s^*_{(B+1)\alpha/2}$ and $s^*_{(B+1)(1-\alpha/2)}$, and the confidence interval is given as

$$\widehat{\theta} - s^*_{(B+1)(1-\alpha/2)} \leq \theta \leq \widehat{\theta} - s^*_{(B+1)\alpha/2} \tag{24}$$

We can rewrite the confidence interval (24) by considering the quantiles of $\widehat{\theta}^*$ instead of the quantiles of the centered version $(\widehat{\theta}^* - \widehat{\theta})$. If r^*_α and s^*_α are respectively the $\alpha-$percentile of the distributions of $\widehat{\theta}^*$ and $(\widehat{\theta}^* - \widehat{\theta})$, then

$$r^*_\alpha = s^*_\alpha + \widehat{\theta} \tag{25}$$

for the simple reason that the two distributions differ only by a shift of $\widehat{\theta}$.

When substituting the relation (25) in (24) the confidence interval based on

the bootstrap distribution is given as

$$2\widehat{\theta} - r^*_{(B+1)(1-\alpha/2)} \le \theta \le 2\widehat{\theta} - r^*_{(B+1)\alpha/2}$$

An important property of the percentile bootstrap interval is that it preserves transformations on the distribution parameters. Supposing that we have adopted a monotone parameter transformation $\widehat{\phi} = m(\widehat{\theta})$ that normalizes the distribution of $\widehat{\theta}$, that is $\widehat{\phi}$ has a Gaussian distribution with mean ϕ and variance c^2, then the percentile interval for θ is

$$m^{-1}(\widehat{\phi} - cz_{(1-\alpha)}) \le \theta \le m^{-1}(\widehat{\phi} - cz_\alpha)$$

This property can be employed mainly when the distribution of $\widehat{\theta}$ is clearly asymmetric and the standard normal interval would be correct if we adopt a transformation that normalizes the distribution. The percentile method makes this transformation automatically. The main advantage of the method is, therefore, not to demand the explicit knowledge of the transformation that normalizes the distribution of $\widehat{\theta}$. However, for the bootstrap percentile method to be precise it is necessary that the transformation exists, but that does not always occur. On these situations, a correction on the percentile method is proposed in Efron and Tibshirani (1993) and it is named BC_a (bias-corrected and accelerated) method. The BC_a intervals are a substantial improvement over the percentile method in both theory and practice, but a little more effort is required.

The ABC method reduces by a large factor the amount of computation required for BC_a method but we do not present it in this paper, see Efron and Tibshirani (1993) for details.

The BC_a method: The construction of the BC_a method is more complicated than the percentile intervals, but it is almost as easy to use. Let $\widehat{\theta}^{*(\alpha)}$ indicate the 100α-th percentile of B bootstrap replications $\widehat{\theta}^*(1), \widehat{\theta}^*(2), \dots,$ $\widehat{\theta}^*(B)$. The percentile interval $(\widehat{\theta}_{lo}, \widehat{\theta}_{up})$ of intended coverage $1 - 2\alpha$ is obtained directly from theses percentiles. If $B = 2000$ and $\alpha = 0.05$, then the interval $(\widehat{\theta}^{*(0.05)}, \widehat{\theta}^{*(0.95)})$ is the interval extending from the 100th to the 1900th ordered values of the 2000 numbers $\widehat{\theta}^*(b)$.

The BC_a interval endpoints are also given by percentiles of the bootstrap distribution, but they are different. The percentiles used depends on two numbers $\widehat{a}$ and $\widehat{z}_0$, called the accelerated and bias-correction. The BC_a interval of intended coverage $1 - 2\alpha$, is given by

$$(\widehat{\theta}_{lo}, \widehat{\theta}_{up}) = (\widehat{\theta}^{*(\alpha_1)}, \widehat{\theta}^{*(\alpha_2)})$$

where

$$\alpha_1 = \phi\left(\widehat{z}_0 + \frac{\widehat{z}_0 + z^{(\alpha)}}{1 - \widehat{a}(\widehat{z}_0 + z^{(\alpha)})}\right), \alpha_2 = \phi\left(\widehat{z}_0 + \frac{\widehat{z}_0 + z^{(1-\alpha)}}{1 - \widehat{a}(\widehat{z}_0 + z^{(1-\alpha)})}\right)$$

Here $\phi(\cdot)$ is the standard normal cumulative distribution function and $z^{(\alpha)}$ is the 100α-th percentile point of a standard normal distribution. Notice that if $\widehat{a} = \widehat{z}_0 = 0$, then the BC_a interval is the same as the percentile interval. Non-zero values of $\widehat{a}$ and $\widehat{z}_0$ correct certain deficiencies of the standard and percentile methods.

The value of the bias-correction $\widehat{z}_0$ is obtained directly from the proportion of bootstrap replications smaller than the original estimate $\widehat{\theta}$,

$$\widehat{z}_0 = \phi^{-1}\left(\frac{\#\{\widehat{\theta}^*(b) < \widehat{\theta}\}}{B}\right)$$

where $\phi^{-1}(\cdot)$ indicates the inverse function of a standard normal cumulative distribution function. Roughly speaking, $\widehat{z}_0$ measures the median bias of $\widehat{\theta}^*$, that is, the discrepancy between the median of $\widehat{\theta}^*$ and $\widehat{\theta}$, in normal units. (We obtain $\widehat{z}_0 = 0$ if exactly half of the $\widehat{\theta}^*(b)$ values are smaller than or equal to $\widehat{\theta}$).

There are many ways to compute the acceleration parameter $\widehat{a}$. The easiest to explain is given in terms of the jackknife values of a statistic $\widehat{\theta} = s(x)$. Let $x_{(i)}$ be the original sample with the i-th point x_i deleted, let $\widehat{\theta}_{(i)} = s(x_{(i)})$, and define $\widehat{\theta}(\cdot) = \sum_{i=1}^n \widehat{\theta}_{(i)}/n$. A simple expression for the acceleration is

$$\widehat{a} = \frac{\sum_{i=1}^n (\widehat{\theta}(\cdot) - \widehat{\theta}_{(i)})^3}{6\{\sum_{i=1}^n (\widehat{\theta}(\cdot) - \widehat{\theta}_{(i)})^2\}^{3/2}}$$

The quantity $\widehat{a}$ is called acceleration because it refers to the rate of change of the standard error of $\widehat{\theta}$ with respect to the true parameter value θ. The standard normal approximation $\widehat{\theta} \sim N(\theta, se^2)$ assumes that the standard error of $\widehat{\theta}$ is the same for all θ. However, this is often unrealistic and the acceleration constant $\widehat{a}$ corrects this behavior.

The BC_a method has two important theoretical advantages. First of all, it preserves transformations on the distribution parameters, as the percentile method. The second advantage concerns its accuracy. A central $1 - 2\alpha$ confidence interval is supposed to have probability α of not covering the true value of θ from above or below. Approximate confidence intervals can be graded on how accurately they match that probability and the BC_a intervals can be shown to be of second-order accurate, meaning that the errors go to zero at

rate $1/n$ in terms of the sample size n. The standard and percentile methods are only of first-order accurate, meaning that the errors are in an order of magnitude larger.

7   Numerical results

In order to verify the performance of the profile maximum likelihood technique (PML) and also of the bootstrap method, we generated two data sets with the Richards growth model; one with 500 and other with 50 observations. Initially, 5000 values were generated with $q = 0.6$, $K = 9000$, $\rho = 0.7$ and $\tau = 2$. The last 1000 values were separated and then the two samples were selected, avoiding the effect of the population initial size. The sample with 500 observations was called Sample 1 and the other with 50 observations was called Sample 2, see Figures 1 and 2 respectively.
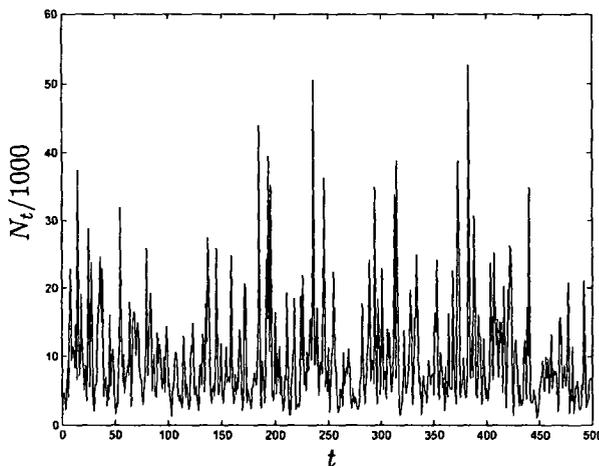
Fig. 1. Sample 1, $n = 500$.

With those samples we calculated the profile maximum likelihood estimates (PMLE) of the model parameters and the corresponding asymptotic intervals. We also calculated the bootstrap intervals of the model parameters with the percentile and BC_a methods. In the figures 3 a) and 3 b) we present the graphs of the log profile likelihood functions in which we drawn the asymptotic confidence intervals for q, for the samples 1 and 2, respectively.

In the Tables 1 and 2 we present the profile maximum likelihood estimates (PMLE) and the comparison of the asymptotic confidence intervals with the percentile and BC_a intervals for the Samples 1 and 2, respectively.

Analysing Tables 1 and 2 we verify that the bootstrap method is more adequate to calculate confidence intervals of the Richards growth model parame-

13
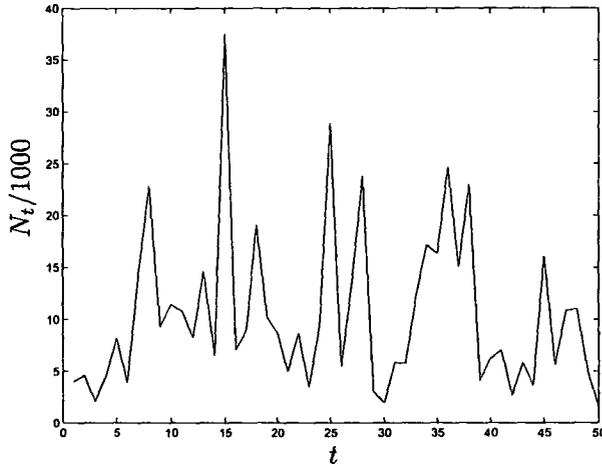
Fig. 2. Sample 2, $n = 50$.

Table 1

Sample 1 $(n = 500)$ $\tau = 2, q = 0.6, K = 9000$ and $\rho = 0.7$.

	PMLE	Asymptotic CI	Percentile CI	BC$_a$ CI
τ	2.2306	[2.2160 ; 2.2452]*	[2 ; 2.5588]	[1.9821 ; 2.5373]
q	0.49	[0.24 ; 0.75]	[0.24 ; 0.72]	[0.2423 ; 0.72]
K	8197.1	[8196.3 ; 8197.9]*	[7367.3 ; 9094.8]	[7382.9 ; 9124.7]
ρ	0.6704	[0.6309 ; 0.7099]	[0.5910 ; 0.7574]	[0.5769 ; 0.7454]

* Intervals that do not contain the true values of the parameters

Table 2

Sample 2 $(n = 50)$ $\tau = 2, q = 0.6, K = 9000$ and $\rho = 0.7$.

	PMLE	Asymptotic CI	Percentile CI	BC$_a$ CI
τ	2.0831	[1.3369 ; 2.8293]	[1.4314 ; 3.4636]	[1.3519 ; 2.9820]
q	0.94	[0 ; 2.11]	[0.1 ; 1]	[0.1 ; 1]
K	10164.8	[10138 ; 10019.1]*	[7119.3 ; 12715.9]	[7737.7 ; 13637.6]
ρ	0.7626	[0.4162 ; 1.1090]	[0.4847 ; 1.1357]	[0.3893 ; 1.0260]

* Interval that does not contain the true value of the parameter

ters. The bootstrap intervals contain the true value of the parameters, a fact that is not verified with the asymptotic intervals, even when dealing with a sample with 500 observations. In the figures 4 and 5 we present the bootstrap estimates histograms. Those histograms approach the PMLE distribution and they show the poor accuracy of the standard technique, esteeming from the variability of the estimates based on the asymptotic theory and from the Fisher information matrix calculus.

14
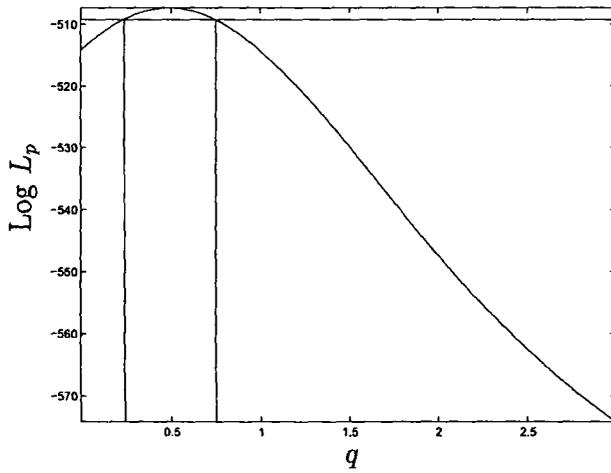
(a)
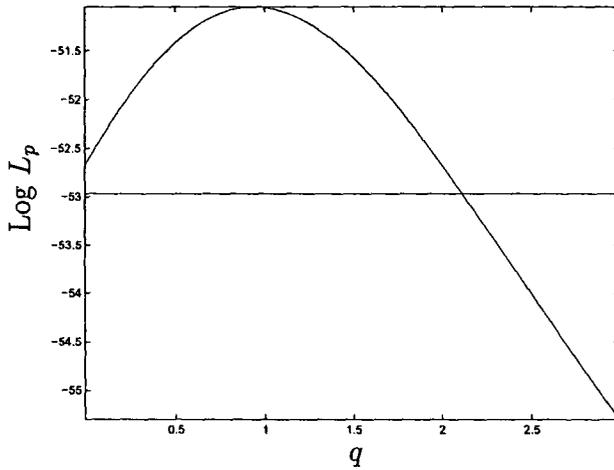
(b)

Fig. 3. Profile Log Likelihood function with CI 95%. a) $(n = 500)$ and $q : (0.24; 0.75)$; b) $(n = 50)$ and $q : (0; 2.11)$.

When $n = 50$, the estimates of the form parameter q indicate values near $q = 1$ with high frequency as shown in the histogram of Figure 5. Even though a robust technique was used for estimate the parameter variability, the estimates lead us to choose erroneously a logistic model $(q = 1)$, whereas the populations are generated by a general Richards model.

15
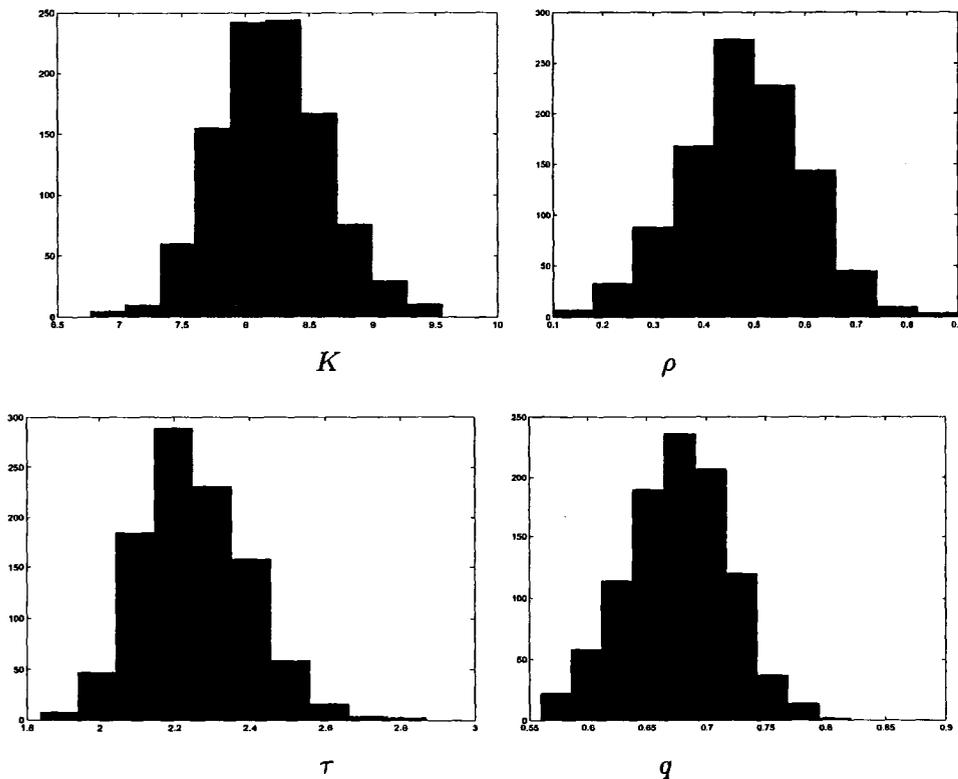
Fig. 4. Bootstrap distributions of K, ρ, τ and q; $n = 500$.

8 Conclusions

The parameter identification problem is fundamental for the studies of the extinction risk, the population management and control, and the population viability. However the identification problem generally faces difficulties connected with small samples, yielding imprecise parameters estimates. Researchers in the field have used the logistic and Gompertz models because of theirs simplicity and this paper shows that the bootstrap technique makes viable the use of more complex models such as the Richards growth model. The aim is to provide better tools for evaluation of the population growth, inferring in a more precise way the population growth rate, the intrinsic growth rate and the carrying capacity, which are important in determining the risk of extinction.

The bootstrap technique was proven accurate but not enough to avoid superestimation of the variability of some parameters when the sample size is small ($n = 50$), and one could erroneously choose a logistic model ($q = 1$). Despite of that, the bootstrap technique should be recommended, since in real world situations one may find a set of data more favourable to the calculus of estimate variability, and fit a more realistic model. This is a more reasonable course of action than simply adopt a logistic model beforehand. Of course a good model

16
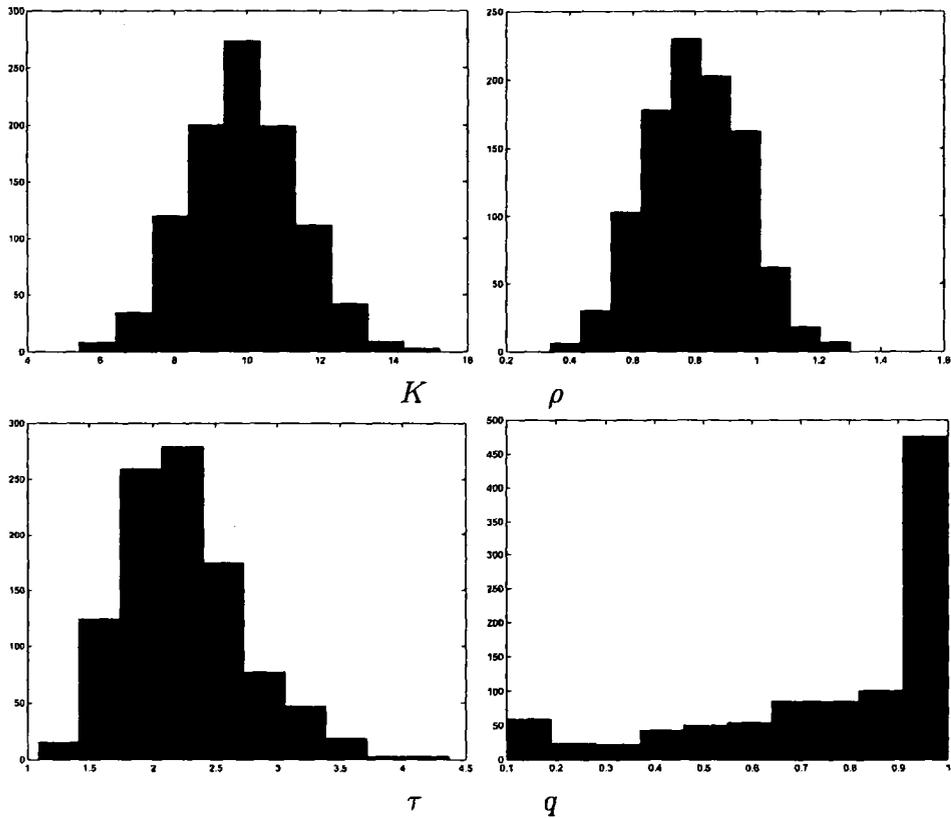
Fig. 5. Bootstrap distributions of K, ρ, τ and q; $n = 50$.

fitting may allow a right decision in an specie preservation policy, or a more accurate prediction of the expected time of population extinction.

Another recommendation for identification of population growth models is to account for specialist information such as biologists, ecologists, and technicians of related fields. These considerations and point of views can be incorporated in the modelling more easily in the Bayesian inference approach, which can be a promising tool in the parameter inference problem of population growth models when only small data samples is available.

Acknowledgement

References

O.E. Barndorff-Nielsen and D.R. Cox, *Inference and Asymptotics* (Chapman & Hall, 1994).

G.E. Box and D.R. Cox, An Analysis of Transformations, *Journal of the Royal Statistic Society*, Serie B, **6** (1964) 211–252.

M.S. Boyce, Population Viability Analysis, *Annu. Rev. Ecol. Syst.* **23** (1992) 481–506.

B. Efron and R.J. Tibshirani, *An introduction to the Bootstrap* (Chapman & Hall, 1993).

H.A. Fitzhugh Jr., Analysis of growth curves and strategies for altering their shapes, *Journal of Animal Science* **42** (1974) 1036–1051.

A. Gelman, J. Carlin, H. Stern and D. Rubin, *Bayesian data Analisys* (Chapman & Hall,1995)

G. Hertzler, J. Herman and R.K. Lindner, Estimating Stochastic Model of Population Dynamics With an Application to Kangaroos, *Natural Resource Modeling* **10** (1997) 303–343.

E.E. Homes, Estimating Risk in Declining Populations with Poor Data, *Proceedings of the National Academy of Sciences or the United States of America* **98** (2001) 5072–5077.

D. Ludwig, Uncertainty and the Assessment of Extinction Probabilities, *Ecological Applications* **6** (1996) 1067–1076.

D. Ludwig, Is It Meaningful to Estimate a Probability of Extinction? *Ecology* **80** (1999) 298–310.

M. Omlin and P. Reichert, A Comparison of Techniques for the Estimation of Model Prediction Uncertainty, *Ecological Modelling* **115** (1999) 45–59.

P. Reichert and M. Omlin, On the Usefulness of Over-parameterized Ecological Model, *Ecological Modelling* **95** (1997) 289–299.

J.F. Richards, A Flexible Growth Function for Empirical Use, *Jor. Exp. Bot.* **10** (1959) 290–.

G.A.F. Seber and C.J. Wild, *Nonlinear Regression* (John Wiley & Sons, 1989).

R. Wehrens, H. Putter and L.M.C. Buydens, *The bootstrap: a tutorial; Chemometrics and Intelligent laboratory systems* (Elsevier, 2000).

NOTAS DO ICMC

SÉRIE ESTATÍSTICA