

APLICAÇÃO DE MODELOS CIRCULARES EM ANÁLISES DA BORDA PARA VALIDAÇÃO EM ESTUDOS DE IDADE DE TUBARÕES

Diego Vicente de Souza FERREIRA¹; Humber Agreli de ANDRADE²; Rosangela Paula Teixeira LESSA²

RESUMO

A análise de validação é fundamental em estudos de idade de peixes. Nas últimas décadas, as pesquisas nesta área foram frequentemente baseadas em análises do incremento marginal em vértebras, e os resultados tem sido interpretados com base em inspeções visuais de gráficos. Interpretações subjetivas podem levar a conclusões equivocadas. Recentemente, modelos estatísticos com a distribuição de probabilidade de Bernoulli e com mistura de distribuições circulares foram propostos como solução objetiva em estudos baseados na análise do tipo de borda. Neste trabalho, diversos destes modelos foram utilizados para analisar dados dos tubarões mako (*Isurus oxyrinchus*) e crocodilo (*Pseudocarcharias kamoharai*). Os resultados sugerem que, em algumas circunstâncias, não se obtém evidências conclusivas. As estimativas dos parâmetros são imprecisas se o tamanho amostral é pequeno. Não há garantias de que uma amostra grande resultará em uma estimação de parâmetros e em uma análise de validação bem sucedida. Para obter evidências conclusivas em análises dos tipos de bordas são necessários modelos adequados e também dados acurados. O uso do Critério de Informação de Akaike pode resultar na seleção de modelos de mistura sobreparametrizados, enquanto que modelos simples são selecionados quando utilizado o Critério de Informação Bayesiano. O uso dos modelos estatísticos circulares é um avanço em relação às abordagens subjetivas, mas mais estudos são necessários sobre distribuições de probabilidade assimétricas e sobre procedimentos para avaliar sobreparametrizações quando utilizados modelos de mistura.

Palavras chave: modelos estatísticos; *Pseudocarcharias kamoharai*; *Isurus oxyrinchus*; critério de informação; anéis etários

ANALYSIS OF EDGE TYPE FOR VALIDATION IN STUDIES OF AGE OF SHARKS: APPLYING CIRCULAR MODELS

ABSTRACT

Validation analysis is important in aging studies of fishes. In the last decades, researchs in this field were often based on marginal increments of vertebrae, and the results have been interpreted based on visual inspection of graphs. Subjective interpretations can result in misleading conclusions. Recently, statistical models with Bernoulli probability distribution and with mixture of circular distributions were proposed as an objective solution in validation studies based on the analysis of the edge type. In this work several of those models were used to analyze data of mako (*Isurus oxyrinchus*) and crocodile (*Pseudocarcharias kamoharai*). In some circumstances results are not conclusive. Parameter estimations are imprecise when the sample size is small. However, there is not guarantee that large samples will result in successful estimation of parameters and validation analysis. Adequate models as well as accurate data are necessary in order to obtain conclusive evidences in the analysis of the edge type. The use of Akaike Information Criterion can result in the selection of overparametrized mixture models, while simple models are selected when using Bayesian Information Criterion. The use of the circular statistical models stand for an advance with respect to the subjective approach, but further studies are necessary concerning asymmetric probability distributions, and procedures to assess overparameterizations when using mixture models.

Keywords: statistical models; *Pseudocarcharias kamoharai*; *Isurus oxyrinchus*; information criterion; aging rings

Artigo Científico: Recebido em 24/01/2014 – Aprovado em 22/01/2015

¹ Departamento de Informática e Estatística, Universidade Federal Rural de Pernambuco (UFRPE). Av. D. Manoel de Medeiros, s/nº – Dois Irmãos – CEP: 52.171-900 – Recife – PE – Brasil. e-mail: diegovicentedesouza@gmail.com (autor correspondente)

² Departamento de Pesca e Aquicultura, UFRPE. Av. D. Manoel de Medeiros, s/nº – Dois Irmãos – CEP: 52171-030 – Recife – PE – Brasil. e-mails: humber.andrade@gmail.com; rptlessa@gmail.com

INTRODUÇÃO

Muitos dos modelos de avaliação de estoques de peixes são baseados em dados estruturados por idade (HILBORN e WALTERS, 1992; QUINN e DERISO, 1999). Estimativas de idades são necessárias para cálculos de parâmetros populacionais importantes, relacionados ao crescimento, à mortalidade e às análises demográficas e, quando pouco acuradas, podem levar à conclusões equivocadas sobre os potenciais produtivos dos recursos pesqueiros, e conseqüentemente, à decisões de manejo inadequadas (GOLDMAN, 2005).

Diferentes abordagens podem ser usadas para obter informações que permitem estimar as idades dos peixes. Entre elas, destaca-se o estudo das estruturas de aposição (*e.g.* vértebras), que pode ser aplicado a elasmobrânquios, ainda que seus esqueletos sejam cartilagosos (CAILLIET *et al.*, 1983). No estudo das estruturas de aposição, uma das alternativas para validar a estimação da idade é a análise da alternância temporal do tipo de borda, a qual é caracterizada como opaca ou translúcida. Essa classificação depende da concentração e natureza do material depositado, o qual influencia na refração, reflexão e absorção de luz (GOLDMAN, 2005; CAILLIET *et al.*, 2006). O material translúcido permite a transmissão ou a passagem de luz, enquanto que o mesmo não ocorre com o opaco. Os termos opaco e translúcido podem ser utilizados independentemente do método de iluminação (CASSELMAN, 1983). A técnica de análise da alternância do tipo de borda está sujeita a imprecisões, mas é frequentemente usada em virtude de seus requisitos modestos, como o baixo custo (CAILLIET *et al.*, 2006).

A composição e característica (opaca ou translúcida) da matriz da estrutura de aposição muda em função de variações de fatores ambientais, que interferem no metabolismo, e também em função do ciclo endógeno (CASSELMAN, 1983). A passagem pelas diferentes etapas do ciclo de vida, ou mesmo a alternância de cenários de maior ou menor temperatura ou disponibilidade de alimento, implicam em alterações da composição, da coloração, da densidade e da velocidade de deposição de material na matriz óssea (teleósteos)

ou cartilaginosa (elasmobrânquios). À medida que o indivíduo envelhece, são então formadas bandas opacas ou translúcidas. Caso sejam encontrados padrões de crescimento, como por exemplo, a formação de um par de bandas por ano, pode-se utilizar a terminologia “anel” ou “annulus” (CAILLIET e GOLDMAN, 2004).

A validação da periodicidade de deposição de uma banda de crescimento como indicativo da idade parece ser especialmente difícil em elasmobrânquios (CAMPANA, 2001; CAILLIET *et al.*, 2006), mas a maioria dos autores assume que há a formação de um anel a cada ciclo anual. Frequentemente, a interpretação dos resultados é baseada na inspeção visual subjetiva de gráficos obtidos na análise da variação do tipo de borda. Poucos pesquisadores têm utilizado critérios objetivos para avaliar a hipótese de formação de um anel anual (CAMPANA *et al.*, 2002; PIERCY *et al.*, 2007). Em alguns casos, são até realizados testes de hipóteses para verificar se há diferenças mensais significativas entre as proporções de determinado tipo de borda, mas não são realizados testes para avaliar a existência ou não da ciclicidade de deposição dos tipos de bordas (OKAMURA e SEMBA, 2009).

Uma abordagem útil para interpretar objetivamente os resultados da análise das variações do tipo de borda é a construção de um modelo de probabilidade para a alternância de bandas (opacas ou translúcidas), baseado na distribuição de Bernoulli. Neste modelo proposto por OKAMURA e SEMBA (2009), a probabilidade (p) de deposição de determinada banda (*e.g.* opaca) em determinado período do ano segue uma função de densidade circular. No artigo dos autores mencionados acima, faz-se uso de uma única distribuição circular ou de uma mistura delas para modelar três diferentes cenários, um em que não há ciclos periódicos de deposição dos anéis, um em que há um ciclo anual, e um em que há dois ciclos anuais. Este último modelo de dois ciclos anuais, adotado por OKAMURA e SEMBA (2009), é simétrico no sentido de que a defasagem entre os dois picos de formação de bandas opacas (ou translúcidas) é de exatamente 6 meses, e de que os picos tem exatamente a mesma intensidade. O desempenho dos modelos adotados foi avaliado a partir de algumas simulações, e é apresentada uma aplicação

somente para um caso específico em que a formação de um anel por ano é bastante evidente, mesmo a partir de inspeção visual de gráficos. Não foram avaliados modelos mais flexíveis e potencialmente mais realistas para cenários de dois ciclos, em que as intensidades e os períodos dos ciclos não são idênticos. Outra questão importante é que não foram também consideradas as incertezas associadas às estimativas dos parâmetros (e.g. intervalos de confiança), o que é fundamental para avaliar se os modelos propostos são adequados.

O objetivo, neste trabalho, foi o de avaliar o potencial de uso dos modelos propostos por OKAMURA e SEMBA (2009) e de adaptações mais flexíveis, em aplicações para casos em que a periodicidade anual na formação da borda é evidente, e também para casos em que os dados são menos informativos. Para esse fim, foram realizadas análises de duas espécies de tubarões (mako - *Isurus oxyrinchus* e crocodilo - *Pseudocarcharias kamoharai*),

para as quais as bases de dados disponíveis apresentaram características distintas.

MATERIAL E MÉTODOS

Dados

Na análise, foram avaliados dados disponíveis para o tubarão mako, capturado no Pacífico Norte, e para o tubarão crocodilo, capturado no oeste do Atlântico Sul. A base de dados referente ao tubarão-crocodilo, cedida pelo Laboratório de Dinâmica de Populações Marinhas do Departamento de Pesca e Aquicultura da Universidade Federal Rural de Pernambuco, contém informação sobre 307 exemplares amostrados na região tropical do sudoeste do Atlântico entre 2003 e 2007 (Tabela 1). Já os dados do tubarão mako são os mesmos que constam no trabalho de OKAMURA e SEMBA (2009), ou seja, 227 tubarões amostrados, entre os anos 1992 e 2005, no Pacífico Norte.

Tabela 1. Número total de vértebras e número e proporção mensal de vértebras com bordas opacas para os tubarões crocodilo (*Pseudocarcharias kamoharai*) e mako (*Isurus oxyrinchus*).

	Meses											
	Jan.	Fev.	Mar.	Abr.	Maió	Jun.	Jul.	Ag.	Set.	Out.	Nov.	Dez.
Tubarão crocodilo												
Total	27	10	20	29	36	31	51	20	30	26	20	7
Bordas opacas	3	1	4	7	4	2	5	2	2	3	3	1
Proporção	0,11	0,10	0,20	0,24	0,11	0,06	0,09	0,10	0,07	0,11	0,15	0,14
Tubarão mako												
Total	31	23	4	12	16	16	22	7	8	45	37	6
Bordas opacas	24	10	1	3	3	2	3	1	1	25	21	4
Proporção	0,77	0,43	0,25	0,25	0,19	0,13	0,14	0,14	0,13	0,56	0,57	0,67

Modelos

Os dados coletados são binários (opaco ou translúcido) e periódicos, portanto, é razoável utilizar modelos que combinam as distribuições de probabilidade de Bernoulli e circulares. O modelo básico é o de OKAMURA e SEMBA (2009), com adaptações quanto à forma com que os parâmetros são tratados. No presente trabalho há uma menor quantidade de restrições e o modelo é mais flexível. A probabilidade de que a borda da vértebra seja opaca é dada por:

$$(1) \quad p = \frac{\gamma}{y(m_i + 1) - y(m_i)} \frac{y(m_i + 1)}{y(m_i)} \int f(u | \theta) du$$

e a função de massa de probabilidade das observações \mathbf{x} dado o vetor de parâmetros θ de interesse é baseada na distribuição de Bernoulli:

$$(2) \quad p(\mathbf{x} | \theta) = \prod_{i=1}^n p^{x_i} (1-p)^{1-x_i}$$

Nas equações acima, o vetor $\mathbf{x} = \{x_{i=1}, \dots, x_{i=n}\}$, em que x_i é igual a 1 para o i -

ésimo peixe se a banda é opaca e 0 se a banda é translúcida. O vetor $\mathbf{m} = \{m_i=1, \dots, m_i=n\}$ denota a covariável em que m_i é o mês em que o i -ésimo exemplar foi coletado. O primeiro dia do mês m_i , expresso em radianos, é $y(m_i)$, ou seja, $y(1) = 0, y(2) = (31/365)2\pi, \dots, y(13) = 2\pi$. γ é o parâmetro de normalização da densidade de probabilidade, usado para garantir que a probabilidade do tipo de borda opaca esteja entre 0 e 1. O parâmetro γ deve satisfazer à restrição:

$$(3) \quad 0 \leq \gamma \leq \left\{ \max \left[\frac{\gamma}{y(m_i+1) - y(m_i)} \frac{y(m_i+1)}{y(m_i)} \int f(u|\boldsymbol{\theta}) du \right] \right\}^{-1}$$

É conveniente que a função de densidade $f(u|\boldsymbol{\theta})$ seja periódica. Uma alternativa é o uso de funções circulares (FISHER, 1995; JAMMALAMADAKA e SENGUPTA, 2001). Neste trabalho, foram consideradas as distribuições de probabilidade circulares de von Mises (VM) e *wrapped* Cauchy (WC) (LEVY, 1939). A distribuição VM é dada por (FISHER, 1995):

$$(4) \quad f(u|\mu, \kappa) = \frac{\exp[\kappa \cos(u - \mu)]}{2\pi I_0(\kappa)}$$

em que: $0 \leq \mu < 2\pi$, $0 \leq \kappa < \infty$, e $I_0(\kappa) = \sum_{i=0}^{\infty} \left\{ \frac{\kappa^{2i}}{2^{2i} (i!)^2} \right\}$ é a função de Bessel modificada de ordem zero. Aproximações para essa função podem ser encontradas em ABRAMOWITZ e STEGUN (1965). Os componentes do vetor de parâmetros $\boldsymbol{\theta}$ são γ , μ que é a média direcional correspondente ao momento (*e.g.* mês) do pico da distribuição de probabilidade de formação de bandas opacas, e κ , que é relacionado à concentração em torno de μ . A média direcional é calculada em radianos, mas a conversão para dias ou meses facilita a interpretação do significado biológico do parâmetro. No caso da VM, quanto menor o κ , mais bem definido é o pico de formação de bandas, com uma proporção de bordas opacas claramente mais elevada em determinado período do ano. Se o valor de κ é elevado, não haveria um pico bem definido e a variação da proporção de bordas opacas é mais difusa. No limite $\kappa \rightarrow \infty$, a equação acima converge para uma uniforme $U(0, 2\pi)$, que poderia ser utilizada para modelar o

cenário em que não há ciclicidade na formação de anéis. Neste caso, o número de parâmetros reduz para $\theta = \gamma$.

A função de densidade WC é obtida se a distribuição Cauchy:

$$(5) \quad f(u|\mu, \sigma) = \frac{1}{\pi} \cdot \frac{\sigma}{\sigma^2 + (u - \mu)^2}$$

é projetada ao redor de um círculo o que resulta em:

$$(6) \quad f(u|\mu, \kappa) = \frac{1}{2\pi} \left[\frac{1 - \kappa^2}{1 + \kappa^2 - 2\kappa \cos(u - \mu)} \right]$$

em que o parâmetro μ é novamente a média direcional tal que em que $0 \leq \mu < 2\pi$ e κ é o parâmetro de concentração que assume um valor entre 0 e 1. Aqui, no caso da WC, a interpretação é que, quanto menor κ , mais difusa é a solução e menor a concentração em torno do pico de formação de bandas opacas. No limite $\kappa \rightarrow 0$, a distribuição converge para a uniforme $U(0, 2\pi)$, a qual representa, então, um caso em que não há uma ciclicidade na formação dos anéis.

Nos modelos indicados acima, pode-se utilizar uma única distribuição circular ou mesmo uma mistura delas, com as devidas atribuições de pesos. Por exemplo, a função

$$(7) \quad f(u|\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) = \omega f(u|\boldsymbol{\theta}_1) + (1 - \omega) f(u|\boldsymbol{\theta}_2)$$

representa uma mistura de duas funções circulares com pesos ω e $1 - \omega$, respectivamente. A formulação pode até ser estendida para mais de duas distribuições. No entanto, neste artigo optou-se pela mistura de no máximo duas distribuições circulares, o que seria suficiente para modelar dois ciclos anuais de formação de anéis. O modelo com a mistura de duas distribuições pode colapsar se $\omega \approx 0$ ou $\omega \approx 1$, ou se $\mu_1 \approx \mu_2$. Se isto ocorre, o modelo de dois ciclos se reduz a um modelo de um ciclo.

Ao final, há três estruturas básicas de modelos (sem ciclo, um ciclo, dois ciclos) e duas funções de densidades $f(u|\boldsymbol{\theta})$ (VM e WC). No trabalho de OKAMURA e SEMBA (2009) foram consideradas as três estruturas mostradas acima, mas foram impostas restrições importantes quanto aos

parâmetros do modelo de dois ciclos (*i.e.* $\omega = 0,5$; $\kappa_1 = \kappa_2$; $\mu_2 = \mu_1 + \pi$). Neste trabalho há maior flexibilidade com a aplicação de doze modelos

plausíveis (Tabela 2). Os três primeiros modelos (A, B, e C) foram os únicos considerados por OKAMURA e SEMBA (2009).

Tabela 2. Modelos considerados. Todos os casos foram avaliados com distribuições circulares de von Mises e *wrapped* Cauchy. Os modelos propostos por OKAMURA e SEMBA (2009) são indicados com asteriscos.

Modelo	Restrições dos Parâmetros					
	γ	μ_1	κ_1	ω	μ_2	κ_2
A*	não	—	—	—	—	—
B*	não	não	não	—	—	—
C*	não	não	não	0,5	$\mu_1 + \pi$	κ_1
D	não	não	não	0,5	$\mu_1 + \pi$	$\kappa_1/2$
E	não	não	não	0,5	$\mu_1 + \pi/2$	κ_1
F	não	não	não	0,5	$\mu_1 + \pi/2$	$\kappa_1/2$
G	não	não	não	0,75	$\mu_1 + \pi$	κ_1
H	não	não	não	0,75	$\mu_1 + \pi$	$\kappa_1/2$
I	não	não	não	0,75	$\mu_1 + \pi/2$	κ_1
J	não	não	não	0,75	$\mu_1 + \pi/2$	$\kappa_1/2$
K	não	não	não	não	não	não

As estimativas dos parâmetros dos modelos foram calculadas numericamente com a minimização de menos logaritmo da verossimilhança com o método simplex (NELDER e MEAD, 1965), implementado com a função `optim()` do programa R (R CORE TEAM, 2011). Para a inferência e cálculo de intervalos de confiança para os parâmetros estimados, foi utilizada a diagonal da matriz Hessiana obtida na otimização, e foi assumida assintoticamente normalidade para a distribuição amostral das estimativas.

Seleção de modelos

Diferentes critérios podem ser usados para comparar modelos concorrentes. O propósito do uso de mais de um critério foi o de obter maiores informações comparativas entre os modelos, uma vez que as diferenças entre as métricas dos diversos critérios podem levar à seleção de modelos distintos. Uma das alternativas é o Critério de Informação de Akaike - *AIC* (AKAIKE, 1974):

$$(8) \quad AIC_i = -2\ln L(\theta | g_i, \text{dados}) + 2k$$

em que k é o número de parâmetros e $\ln L(\theta | g_i, \text{dados})$ é o logaritmo natural da

verossimilhança avaliada no estimador de máxima verossimilhança de θ , condicionada aos dados e à estrutura do i -ésimo modelo g . Neste estudo, o modelo mais complexo (K na Tabela 2) tem o vetor $\theta = \{\gamma, \mu_1, \kappa_1, \omega, \mu_2, \kappa_2\}$ de seis parâmetros a serem estimados. Os demais casos correspondem a reduções deste modelo.

No *AIC* é avaliado o balanço entre o viés e a variância em virtude da inclusão de uma maior ou menor quantidade de parâmetros. Quanto menor o *AIC*, melhor o modelo quanto ao balanço mencionado acima. Diferenças de valores de *AIC* maiores que 2 usualmente indicam que um dos modelos é superior ao outro (BURNHAM e ANDERSON, 2002).

A verossimilhança do modelo g_i , ou seja, $L(g_i | \text{dados})$, é proporcional à diferença entre o *AIC* do i -ésimo modelo e o *AIC* mínimo encontrado nos cálculos para os demais modelos avaliados (AKAIKE, 1983):

$$(9) \quad L(g_i | \text{dados}) \propto \exp\{-0,5[AIC_i - \min(AIC)]\}$$

Já o peso de evidência em favor do i -ésimo modelo é (AKAIKE, 1983):

$$(10) \quad WAIC_i = \frac{L(g_i | \text{dados})}{\sum L(g_i | \text{dados})}$$

Assim, depois de calculados, os pesos podem ser utilizados para criar um ranking dos modelos mais adequados entre os avaliados.

Uma adaptação do critério de informação de Akaike, proposta por SUGIURA (1978), é o *AIC* de segunda ordem (*AICc*), que é usualmente recomendado se o tamanho da amostra não é muito maior que o número de parâmetros. BURNHAM e ANDERSON (2002) sugerem o uso do *AICc* ao invés do *AIC* quando a relação entre tamanho da amostra (n) e o número de parâmetros (k) é menor que 40. O *AICc* é dado por:

$$(11) \quad AICc_i = AIC_i + \frac{2k(k+1)}{n-k-1}$$

O *AICc* também pode ser usado para o cálculo dos pesos de evidências favoráveis aos diversos modelos. Uma outra alternativa é o Critério de Informação de Bayes - *BIC* (SCHWARZ, 1978):

$$(12) \quad BIC_i = -2\log L(\theta | g_i, \text{dados}) + k \log(n)$$

Neste caso, pode ser calculada uma probabilidade posterior (*PBIC*) para modelos concorrentes na forma:

$$(13) \quad PBIC = \frac{\exp(-0,5\Delta_i)}{\sum_{i=1}^R \exp(-0,5\Delta_i)}$$

em que $\Delta_i = BIC_i - \min(BIC)$. Estas probabilidades posteriores podem também ser utilizadas para construir um ranking dos modelos concorrentes.

Comparações baseadas nos critérios e medidas quantitativas mencionadas acima são úteis para a seleção entre os doze modelos considerados (Tabela 2). No entanto, é importante destacar que a validade das comparações dos demais modelos com o modelo complexo de mistura de dois ciclos sem restrições (K), está condicionada a que não ocorra o colapso deste modelo, de maior complexidade, para um modelo simples de um ciclo.

RESULTADOS

Tubarão mako

Os valores do *AICc* dos diferentes modelos ajustados aos dados do tubarão mako foram intermediários entre os do *AIC* e do *BIC* (Tabela 3). Comparando-se os valores de *AIC*, *AICc* e *BIC* obtidos com as duas distribuições (VM e WC), verificou-se que, na maioria dos casos, os valores dos critérios de informação foram menores para a VM. Isto é um indício de que essa distribuição é, no geral, mais adequada para modelar os dados do tubarão mako.

Tabela 3. Critérios de informação de Akaike (*AIC*), de Akaike corrigido (*AICc*) e Bayesiano (*BIC*) aplicados para o caso do tubarão mako (*Isurus oxyrinchus*). Pesos de evidência a favor de cada modelo (Mod) calculados com base no *AIC* (PA) e no *AICc* (PAC). Probabilidade posterior (PB) calculada com base no *BIC*. As distribuições circulares consideradas são a de von Mises (VM) e a *wrapped* Cauchy (WC).

Mod	VM			WC			VM			WC		
	<i>AIC</i>	<i>AICc</i>	<i>BIC</i>	<i>AIC</i>	<i>AICc</i>	<i>BIC</i>	PA	PAC	PB	PA	PAC	PB
A	312,44	312,46	315,87	312,44	312,46	315,87	0,000	0,000	0,000	0,000	0,000	0,000
B*	271,89	271,99	282,16	273,79	273,90	284,07	0,883	0,888	0,923	0,341	0,384	0,953
C	316,24	316,35	326,51	316,27	316,38	326,55	0,000	0,000	0,000	0,000	0,000	0,000
D	297,06	297,17	307,34	298,84	298,95	309,12	0,000	0,000	0,000	0,000	0,000	0,000
E	285,23	285,34	295,50	283,04	283,14	293,31	0,001	0,001	0,001	0,003	0,004	0,009
F	292,04	292,15	302,32	292,88	292,99	303,15	0,000	0,000	0,000	0,000	0,000	0,000
G	283,20	283,31	293,47	289,30	289,40	299,57	0,003	0,003	0,003	0,000	0,000	0,000
H	277,92	278,03	288,20	282,37	282,48	292,65	0,043	0,043	0,045	0,005	0,005	0,013
I	283,46	283,57	293,73	285,39	285,50	295,66	0,003	0,003	0,003	0,001	0,001	0,003
J	279,15	279,26	289,43	281,79	281,90	292,90	0,023	0,023	0,025	0,006	0,007	0,012
K*	277,89	278,27	298,44	272,52	273,01	293,18	0,044	0,039	0,000	0,644	0,599	0,010

* Modelos com maiores evidências a seu favor.

Entre os três modelos (A, B e C) avaliados por OKAMURA e SEMBA (2009), o modelo B (um ciclo) foi sempre o selecionado, independente do critério de informação e da distribuição de probabilidade utilizada. No entanto, quando considerados também os novos modelos, o de mistura com dois ciclos sem restrições (K), apresentou pesos de evidências superiores aos dos demais modelos, quando se fez uso da distribuição WC e do AICs. Porém, o modelo K

colapsou para um ciclo, pois o valor da estimativa pontual do parâmetro ω foi de aproximadamente zero, e o amplo intervalo de confiança incluiu valores negativos impróprios (Tabela 4). Isto indica que o modelo está sobreparametrizado, com mais parâmetros que o necessário. Assim, ainda que os pesos de evidências favoráveis ao modelo K sejam elevados, a hipótese de existência de um ciclo aparece claramente como a mais plausível para o mako.

Tabela 4. Estimativas dos parâmetros e respectivos intervalos de confiança (95%) dos modelos selecionados para o tubarão mako (*Isurus oxyrinchus*). γ é o parâmetro de normalização, μ_1 e μ_2 são as médias direcionais em radianos, e κ_1 e κ_2 são os parâmetros de concentração. Modelos (Mod): simples de um ciclo (B), e de mistura sem restrições (K); Distribuições (Dist): von Mises (VM) e *wrapped* Cauchy (WC).

Mod	Dist	γ	μ_1	κ_1	ω	μ_2	κ_2
B	VM	2,38±0,56	6,10±1,36	0,88±0,35	---	---	---
B	WC	2,45±0,89	6,11±1,24	0,36±0,42	---	---	---
K	WC	2,27±56,95	5,19±0,11	0,95±11,14	0,09±0,96	6,28±0,65	0,44±0,49

As variações das proporções de bandas opacas e os ajustes dos modelos com maiores pesos de evidência a seu favor são mostrados na Figura 1 A. Note que os ajustes do modelo B (um ciclo) se sobrepõe bastante ao ajuste do modelo K (dois ciclos sem restrições), o que ilustra o colapso deste último. Foi também visível a alternância na predominância de exemplares com bordas de vértebras translúcidas ou opacas. As proporções de bordas opacas foram elevadas

nos meses de dezembro e janeiro e reduzidas no meio do ano, especialmente entre junho e setembro. A amplitude de variação das proporções de bandas opacas foi elevada, sendo que foi atingido um mínimo de 0,13 (junho e setembro) e um máximo de 0,77 (janeiro). Portanto, a base de dados para o mako se caracterizou por apresentar um ciclo anual bastante evidente, mesmo a partir de uma inspeção visual gráfica.

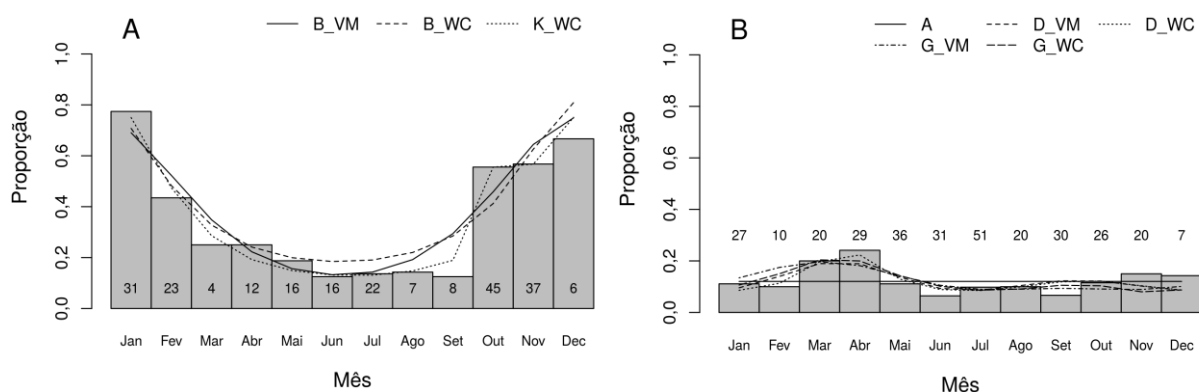


Figura 1. Proporção de bordas opacas (barras) e ajustes dos modelos selecionados (linhas) para os tubarões mako (*Isurus oxyrinchus*) (A) e crocodilo (*Pseudocarcharias kamoharai*) (B). Os tamanhos amostrais são apresentados com numerais menores dentro ou acima das barras nos painéis. As siglas VM e WC referem-se às distribuições de von Mises e *wrapped* Cauchy, respectivamente.

Muitas das estimativas dos parâmetros dos modelos selecionados (B e K) foram imprecisas, como evidenciado pelas grandes amplitudes dos intervalos de confiança, com margens de erro que chegaram a superar as estimativas pontuais (Tabela 4). A imprecisão é especialmente grande nas estimativas obtidas com o modelo de mistura de duas distribuições circulares, que, como mencionado anteriormente, também mostrou claros sinais de colapso e sobreparametrização.

A interpretação das estimativas das médias direcionais (μ), que tem significado biológico, é facilitada se as estimativas são convertidas de radianos para meses ou dias. Assim, as estimativas do parâmetro μ_1 para o modelo com um ciclo (B) para o mako, sugerem um pico de

probabilidade de formação de bandas opacas nas imediações da transição entre dezembro e janeiro (6,10 radianos, que corresponde ao $6,10 \times 365 / 2\pi \approx 354^\circ$ dia do ano), com intervalo de confiança de cerca de 5 meses (157 dias) para o ajuste com a distribuição VM e de 144 dias para o ajuste com a distribuição WC. Os valores de κ também indicam uma concentração em torno do pico do final do ano.

Tubarão crocodilo

Os critérios de informação calculados para os diferentes modelos são mostrados na Tabela 5. Os valores de *AIC*, *AICc* e *BIC* obtidos com as distribuições WC e VM, apresentaram diferenças pequenas, indicando que não houve uma distribuição claramente superior à outra.

Tabela 5. Critérios de informação de Akaike (*AIC*), de Akaike corrigido (*AICc*) e Bayesiano (*BIC*) aplicados para o caso do tubarão crocodilo (*Pseudocarcharias kamoharai*). Pesos de evidência a favor de cada modelo (Mod), calculados com base no *AIC* (PA) e no *AICc* (PAC). Probabilidade posterior (PB), calculado com base no *BIC*. As distribuições circulares consideradas são a de von Mises (VM) e a *wrapped* Cauchy (WC).

Mod	VM			WC			VM			WC		
	<i>AIC</i>	<i>AICc</i>	<i>BIC</i>	<i>AIC</i>	<i>AICc</i>	<i>BIC</i>	PA	PAC	PB	PA	PAC	PB
A*	227,93	227,94	231,66	227,93	227,94	231,66	0,153	0,158	0,885	0,134	0,138	0,868
B	228,96	229,04	240,14	228,82	228,90	240,14	0,091	0,091	0,013	0,086	0,085	0,012
C	230,08	230,16	241,26	229,86	229,94	241,04	0,052	0,052	0,007	0,051	0,051	0,008
D*	228,09	228,17	239,27	227,32	227,40	238,50	0,141	0,141	0,020	0,182	0,181	0,028
E	228,97	229,05	240,15	228,78	228,86	239,96	0,091	0,091	0,013	0,087	0,087	0,014
F	229,35	229,43	240,53	229,74	229,82	240,92	0,075	0,075	0,010	0,054	0,054	0,008
G*	228,58	228,65	239,76	228,07	228,15	239,25	0,111	0,111	0,015	0,125	0,125	0,019
H	228,83	228,91	240,01	228,20	228,28	239,38	0,098	0,097	0,014	0,117	0,117	0,018
I	229,17	229,25	240,35	229,42	229,50	240,60	0,082	0,082	0,011	0,072	0,072	0,011
J	229,10	229,18	240,28	229,10	229,18	240,28	0,085	0,085	0,012	0,075	0,074	0,012
K	232,02	232,30	254,38	231,99	232,27	254,35	0,020	0,018	0,000	0,018	0,016	0,000

* Modelos com maiores evidências a seu favor.

A seleção de modelos, no caso do tubarão crocodilo, foi bem mais complexa que no caso do tubarão mako, especialmente quando utilizados os critérios *AIC* e *AICc* (Tabela 5). Com estes critérios, não houve claramente um modelo mais adequado, e pelo menos três modelos, A, D e G, apareceram como razoáveis. É importante lembrar que estes modelos concorrentes representam hipóteses conflitantes. O modelo A aponta para a inexistência de um ciclo de deposição de anéis, enquanto que os modelos D e G apontam para a

existência de dois ciclos. A incerteza se torna ainda maior se consideramos a possibilidade de que sejam usadas distribuições de probabilidade circulares diferentes. Se utilizado o *AIC* com a VM, o modelo A aparece como levemente superior ao D, mas se utilizamos o *AIC* com a WC, o modelo D passa a figurar como mais adequado que o A. Se utilizado o *BIC* como critério, a confusão quanto à seleção de modelos deixa de existir. Neste caso, o modelo mais simples (A) (sem ciclo) aparece claramente como o mais

adequado. Todos os demais modelos mostram desempenhos inferiores.

O modelo sem ciclo (A) apresenta, por definição, uma proporção fixa de 0,12 para qualquer mês, independentemente da distribuição de probabilidade considerada. Já os modelos de mistura selecionados (D e G) ajustaram melhor que o modelo sem ciclo, mas ainda assim, houve falhas evidentes, pois as proporções de bordas opacas foram subestimadas em novembro e dezembro e superestimadas em fevereiro, junho e setembro (Figura 1B).

No caso do tubarão crocodilo, foram detectadas bordas translúcidas na grande maioria dos tubarões (>80%) em todos os meses do ano. Não houve uma situação em que predominaram exemplares com vértebras com bordas opacas. A amplitude das oscilações das proporções de bandas opacas foi muito pequena (0,18), com mínimo de 0,06 e máximo de 0,24. Portanto, a

mera inspeção visual de gráficos das proporções dos tipos de borda não foi esclarecedora quanto à periodicidade da formação de anéis (Figura 1 B).

As estimativas pontuais e os intervalos de confiança dos modelos selecionados para o tubarão crocodilo são mostrados na Tabela 6. As estimativas foram relativamente precisas daquelas obtidas para o parâmetro de concentração (κ_1) se utilizada a distribuição WC. As estimativas de μ_1 para os modelos de mistura sugerem um pico de probabilidade de formação de bandas opacas na transição dos meses de fevereiro e março, com intervalo de confiança de cerca de 1 mês. Os modelos de mistura selecionados foram restritos a que $\mu_2 = \mu_1 + \pi$, o que implica em um segundo pico com defasagem de 6 meses, localizado portanto, entre setembro e outubro. Este segundo pico teria importância secundária, com parâmetro de concentração reduzido (modelo D) ou peso reduzido (modelo G).

Tabela 6. Estimativas dos parâmetros e respectivos intervalos de confiança (95%) dos modelos selecionados para o tubarão crocodilo (*Pseudocarcharias kamoharui*). γ é o parâmetro de normalização, μ_1 e μ_2 são as médias direcionais em radianos, e κ_1 e κ_2 são os parâmetros de concentração. Modelos (Mod): sem ciclo (A), de mistura com restrições (D e G); Distribuições (Dist): von Mises (VM) e *wrapped* Cauchy (WC).

Mod	Dist	γ	μ_1	κ_1	ω	μ_2	κ_2
A	VM - WC	0,757 ± 0,143	---	---	---	---	---
D	VM	0,768 ± 0,150	1,540 ± 0,419	1,701 ± 0,445	0,5	$\mu_1 + \pi$	$\kappa_1/2$
D	WC	0,760 ± 0,305	1,620 ± 0,331	0,563 ± 0,642	0,5	$\mu_1 + \pi$	$\kappa_1/2$
G	VM	0,778 ± 0,178	1,343 ± 0,582	0,904 ± 0,561	0,75	$\mu_1 + \pi$	κ_1
G	WC	0,767 ± 0,186	1,518 ± 0,445	0,380 ± 0,574	0,75	$\mu_1 + \pi$	κ_1

De maneira geral, os resultados para o tubarão crocodilo indicaram que não há um modelo claramente superior aos demais, e que os modelos selecionados representam hipóteses conflitantes (sem ciclo ou dois ciclos), sendo que a inexistência de um ciclo de deposição dos anéis é uma hipótese plausível.

DISCUSSÃO

Na maioria das análises realizadas para o tubarão mako, os cálculos dos critérios de informação foram levemente menores para a VM em comparações de modelos com estruturas similares e mesmo número de parâmetros. Já no

caso do tubarão crocodilo, critérios de informação foram um pouco menores para a WC. No entanto, a escolha pela VM ou pela WC não interferiu muito nos resultados, visto que os modelos selecionados foram, nos dois casos, quase sempre os mesmos. Recomenda-se, para simplificação, o uso de somente uma delas.

No caso do tubarão mako, as evidências favoráveis a um ciclo foram muito claras e, mesmo por inspeção visual, dificilmente se chegaria a outra conclusão. No entanto, é importante ressaltar o caso particular em que o uso do AIC resultou na seleção do modelo de mistura sem restrições, flexível o suficiente para

descrever dois ciclos, mas que colapsa para um ciclo. Assim, o uso do *AIC* leva, neste caso, à seleção de um modelo sobreparametrizado.

Tanto o uso do *AIC* quanto do *BIC* acabaram por levar à seleção de modelos favoráveis à mesma hipótese no caso do tubarão mako. No entanto, isto não ocorreu na análise do tubarão crocodilo. Quando os resultados obtidos não são concordantes, evidencia-se que a principal diferença é que o *AIC* tende a indicar modelos mais complexos como adequados, enquanto que o *BIC* frequentemente resulta em pesos de evidência elevados para modelos mais simples. Assim, o modelo sem ciclo (A) frequentemente apareceu entre os mais adequados, quando não o melhor, se utilizado o *BIC*. OKAMURA e SEMBA (2009) optaram pelo uso do *AIC* por conta da colocação de BURNHAM e ANDERSON (2002) de que o *BIC* teria melhor desempenho assintoticamente com tamanhos amostrais extremamente elevados e que, para casos com tamanhos amostrais menores, o *AIC* e suas adaptações (*e.g.* *AICc*) seriam mais indicados. No entanto, é importante notar que a colocação de BURNHAM e ANDERSON (2002) sobre a superioridade do *AIC* está baseada em simulações que não contemplaram o problema de mistura de distribuições em que modelos podem colapsar.

Se o *AIC* é tomado como superior ao *BIC*, como preconiza BURNHAM e ANDERSON (2002), a conclusão seria de há uma grande dúvida para o tubarão crocodilo. São igualmente plausíveis as hipóteses extremas de que não exista ciclo algum e de que existam dois ciclos defasados em seis meses. Assim, é importante destacar que o uso da técnica de modelagem proposta por OKAMURA e SEMBA (2009) é de grande utilidade e elimina a subjetividade das inspeções visuais, mas não resulta necessariamente no esclarecimento sobre a validação na análise da alternância dos tipos de borda. Se é difícil discernir os tipos das bordas, as incertezas são grandes e as dúvidas podem persistir, mesmo após a execução de análises sofisticadas.

O processo de deposição de anéis etários está sujeito a variáveis abióticas e bióticas (CAMPANA *et al.*, 2002), relacionadas à reprodução, à frequência de ocorrência dos partos e a fatores endógenos (BROTHERS, 1983; LESSA

et al., 2006), que podem levar a padrões temporais complexos. Isto implica que, em alguns casos, pode não haver de fato um padrão claro, ou que poderia não haver necessariamente uma simetria perfeita entre os picos e cavas em casos com um ou dois ciclos de deposição de anéis. Para avaliar o caso com dois ciclos, OKAMURA e SEMBA (2009) consideraram somente o modelo de simetria perfeita (defasagem de picos de seis meses $\mu_2 = \mu_1 + \pi$ e igualdade de pesos e concentrações dos picos, $\omega = 0,5$ e $\kappa_2 = \kappa_1$) (modelo C). No presente trabalho, o modelo de simetria perfeita não constou entre os de dois ciclos selecionados para o tubarão crocodilo (modelos D e G). Ou seja, modelos assimétricos podem surgir como hipóteses mais razoáveis, e devem, portanto, ser explorados em estudos de validação.

Uma das características dos dados do tubarão crocodilo, que chamou a atenção, foi a grande predominância de bordas translúcidas em praticamente todos os meses. Ou seja, não se observa uma alternância da predominância de tipos de borda. Ainda que este resultado diferenciado possa, de fato, estar refletindo um fenômeno natural, a hipótese de que há um viés com subestimativa do número de vértebras com bordas opacas, não pode ser descartada. Quando não se dispõe de informações confiáveis sobre o tipo de borda, e há dados sobre os incrementos marginais, pode-se, como alternativa, usar os modelos circulares também para a análise do incremento marginal relativo (OKAMURA *et al.*, 2013).

A amostra para o tubarão crocodilo ($n = 307$) foi relativamente grande se comparada à do tubarão mako ($n = 227$), porém, no caso do tubarão crocodilo, não se chegou a uma conclusão após a análise dos tipos de bordas. Assim, evidencia-se que tamanhos amostrais grandes não garantem a obtenção de sucesso na análise. A dificuldade aumenta com tamanhos amostrais pequenos. A título de ilustração, os modelos aqui apresentados foram também ajustados a uma amostra de 25 exemplares do tubarão leopardo (*Triakis semifasciata*), extraída do artigo de SMITH (1984), amostra essa que também foi analisada por OKAMURA e SEMBA (2009). Os autores supracitados se baseiam na seleção de modelos

com critérios de informação para concluir que na pequena amostra do tubarão leopardo, há suporte para a hipótese de formação de um par de bandas por ano. No entanto, OKAMURA e SEMBA (2009) não fizeram inferências estatísticas e não avaliaram a precisão na estimativas dos parâmetros dos modelos. Como a amostra é pequena, os intervalos de confiança foram muito amplos, incluindo valores impróprios (e.g. $\kappa < 0$), o que desqualifica a análise como conclusiva. No nosso entender, a identificação de um bom modelo não deve ser baseada unicamente em critérios de informação. Incertezas quanto às estimativas e, quando aplicável, a coerência conceitual dentro do contexto do problema (e.g. significado biológico), devem também ser levadas em conta.

CONCLUSÕES

O uso de modelos estatísticos circulares nem sempre é o suficiente para esclarecer as questões de validação da periodicidade de deposição de anéis etários em tubarões a partir do estudo das variações do tipo de borda. O sucesso da abordagem depende fundamentalmente da qualidade dos dados disponíveis. É importante considerar modelos múltiplos que representem ao máximo as diversas hipóteses plausíveis quanto à ciclicidade de deposição de anéis etários. O Critério de Informação de Akaike leva à seleção de modelos mais complexos que o Critério de Informação Bayesiano. Além dos critérios de informação, deve-se levar em conta a precisão e a utilidade das estimativas dos parâmetros no processo de seleção de modelos.

AGRADECIMENTOS

Somos gratos a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão de uma bolsa de estudos para o primeiro autor. Agradecemos também à equipe do Laboratório de Dinâmica de Populações Marinhas (DIMAR) da Universidade Federal Rural de Pernambuco (UFRPE) por terem gentilmente cedido os dados de parte de suas leituras de idade para o tubarão crocodilo, analisados neste trabalho. Estamos também em débito com dois revisores anônimos, que em muito contribuíram para enriquecer o trabalho.

REFERÊNCIAS

- ABRAMOWITZ, M. e STEGUN, I.A. 1965 Handbook of mathematical functions: with formulas, graphs, and mathematical tables. Volume 55. 9ª ed. New York: Dover Publications. 1046p.
- AKAIKE, H. 1974 A new look at the statistical model identification. *Automatic Control, IEEE Transactions*, 19: 716-732.
- AKAIKE, H. 1983 Information measures and model selection. *Bulletin of the International Statistical Institute*, 50: 277-291.
- BROTHERS, E.B. 1983 Summary of round table discussions on age validation. In: PRINCE, E.D. e PULOS, L.M. (eds) *Proceedings of the international workshop on age determination of oceanic pelagic fishes: tunas, billfishes, and sharks*. NOAA Technical Report NMFS8. p.35-44.
- BURNHAM, K. e ANDERSON, D. 2002 Model selection and multimodel inference: a practical information-theoretic approach. 2ª edição. New York: Springer-Verlag. 488p.
- CAILLIET, G.M. e GOLDMAN, K.J. 2004 Age determination and validation in chondrichthyan fishes. In: CARRIER J.; MUSICK, J.A.; HEITHANS, M.R. (eds) *Biology of sharks and their relatives*. Boca Raton: CRC Press. p.399-447.
- CAILLIET, G.M.; MARTIN, L.K.; KUSHER, D.; WOLF, P.; WELDON, B.A. 1983 Techniques for enhancing vertebral bands in age estimation of California elasmobranchs. In: PRINCE, E.D. e PULOS, L.M. (eds) *Proceedings of the international workshop on age determination of oceanic pelagic fishes: tunas, billfishes and sharks*. NOAA Technical Report NMFS8. p.157-165.
- CAILLIET, G.M.; SMITH, W.D.; MOLLET, H.F.; GOLDMAN, K.J. 2006 Age and growth studies of chondrichthyan fishes: the need for consistency in terminology, verification, validation, and growth function fitting. *Environmental Biology of Fishes*, 77: 211-228.
- CAMPANA, S.E. 2001 Accuracy, precision and quality control in age determination, including a review of the use and abuse of age validation methods. *Journal of Fish Biology*, 59: 197-242.
- CAMPANA, S.E.; NATANSON, L.J.; MYKLEVOLL, S. 2002 Bomb dating and age determination of large pelagic sharks. *Canadian Journal of Fisheries and Aquatic Sciences*, 59: 450-455.

- CASSELMAN, J.M. 1983 Age and growth assessment of fish from their calcified structures - techniques and tools. In: PRINCE, E.D. e PULOS, L.M. (ed.) *Proceedings of the international workshop on age determination of oceanic pelagic fishes: tunas, billfishes, and sharks*. NOAA Technical Report NMFS8. p.1-18.
- FISHER, N.I. 1995 *Statistical Analysis of Circular Data*. Australia: Cambridge University Press. 514p.
- GOLDMAN, K. 2005 Age and growth of elasmobranch fishes. In: MUSICK, J.A. e BONFIL, R. *Management techniques for elasmobranch fisheries*. Edited by Food and Agriculture Organization of the United Nations (FAO). *FAO Fisheries Technical Paper*. p.76-102.
- HILBORN, R. e WALTERS, C.J. 1992 *Quantitative fisheries stock assessment*. New York: Chapman and Hall. 570p.
- JAMMALAMADAKA, S.R. e SENGUPTA, A. 2001 *Topics in circular statistics* (Series Multivariate analysis: vol. 5). Singapore: World Scientific Press. 335p.
- LESSA, R.; SANTANA, F.M.; DUARTE-NETO, P. 2006 A critical appraisal of marginal increment analysis for assessing temporal periodicity in band formation among tropical sharks. *Environmental Biology of Fishes*, 77: 309-315.
- LEVY, P. 1939 L'addition des variables aléatoires définies sur une circonférence. *Bulletin de la Société Mathématique*, 67: 1-41.
- NELDER, J.A. e MEAD, R. 1965 A simplex method for function minimization. *The Computer Journal*, 7: 308-313.
- OKAMURA, H. e SEMBA, Y. 2009 Novel statistical method for validating the periodicity of vertebral growth band formation in elasmobranch fishes. *Canadian Journal of Fisheries and Aquatic Sciences*, 66: 771-780.
- OKAMURA, H.; PUNT, A.E.; SEMBA, Y.; ICHINOKAWA, M. 2013 Marginal increment analysis: a new statistical approach of testing for temporal periodicity in fish age verification. *Journal of Fish Biology*, 82: 1239-1249.
- PIERCY, A.; CARLSON, J.; SULIKOWSKI, J.; BURGESS, G. 2007 Age and growth of the scalloped hammerhead shark, *Sphyrna lewini*, in the north-west Atlantic Ocean and Gulf of Mexico. *Marine Freshwater Research*, 58: 34-40.
- QUINN, T. J. e DERISO, R.B. 1999 *Quantitative fish dynamics*. New York: Oxford University Press. 542p.
- R CORE TEAM. 2011 *R: A language and environment for statistical computing*. R Foundation for Statistical Computing 1, 409. Disponível em: <<http://www.r-project.org>> Acesso em: 26 Ago. 2011.
- SCHWARZ, G. 1978 Estimating the dimension of a model. *Annals of Statistics*, 6: 461-464.
- SMITH, S.E. 1984 Timing of vertebral band deposition in tetracycline injected leopard sharks. *Transactions of American Fisheries Society*, 113: 308-313.
- SUGIURA, N. 1978 Further analysis of the data by Akaike's information criterion and the finite corrections. *Communications in Statistics Theory and Methods*, 7: 13-26.