UTILIZAÇÃO DA REGRESSÃO ISOTÔNICA EM ESTUDOS DE CURVAS DE CRESCIMENTO

Adriano RODRIGUES[1]
Lucas Monteiro CHAVES[1]
Fabyano Fonseca e SILVA[2]
Walmes Marques ZEVIANI[3]

▪ **RESUMO**: Modelos de regressão não-linear têm se mostrado adequados para descrever curvas de crescimento de animais domésticos de interesse Zootécnico, pois apresentam parâmetros que podem ser interpretados biologicamente. Estes modelos são ajustados a dados de peso-idade por meio de algoritmos iterativos, como o de Gauss-Newton. Um problema frequentemente relatado é não convergência deste algoritmo na presença de oscilações na trajetória esperada da curva, caracterizadas muitas vezes por uma perda de peso abrupta dos animais decorrente da influencia de efeitos ambientais como pouca disponibilidade de forragens e/ou presença de enfermidades. Assim, faz-se necessário o desenvolvimento de procedimentos de estimação que contemplem tal fato, e que de alguma forma considerem a naturalidade da resposta esperada nos experimentos em questão. O objetivo do presente trabalho foi propor uma metodologia de transformação de dados, via análise de regressão isotônica, para estudos de curvas de crescimento cujos dados apresentam distúrbios caracterizados por decréscimos de pesos em determinadas faixas de idades. Além de investigar a eficiência da metodologia baseada em regressão isotônica em relação ao aumento da convergência e da qualidade do ajuste do modelo, objetivou-se também propor um procedimento iterativo de isotonização cujo intuito foi obter uma transformação eficiente para os dados. Todas as metodologias mencionadas foram avaliadas por meio de um estudo de simulação Monte Carlo, possibilitando verificar que as metodologias de isotonização adotados resultaram em maiores porcentagens de convergência e menores erros quadráticos médios (EQM) para os parâmetros dos modelos Logísticos, Von Bertalanffy e Gompertz.

▪ PALAVRAS-CHAVE: Modelos logístico; Von Bertalanffy e Gompertz; transformação de dados; isotonização.

1 Introdução

O estudo do crescimento de animais é extremamente importante para a determinação de técnicas adequadas de manejo, bem como para o melhoramento genético, uma vez que possibilita a identificação de indivíduos que crescem com maior eficiência, e tal

[1] Universidade Federal de Lavras - UFLA, Departamento de Ciências Exatas, Caixa Postal 30, CEP 37260-000, Perdões, MG, Brasil. E-mail: *adrianorodrigues@unis.edu.br* / *lucas@ufla.br*
[2] Universidade Federal de Viçosa – UFV, Departamento de Estatística, CEP 36570-000, Viçosa, MG, Brasil. E-mail: *fabyano@dpi.ufv.br*
[3] Universidade Federal do Paraná – UFPR, Laboratorio de Estatística e Geoinformação – LEG, Departamento de Estatística, CEP 81531-990, Curitiba, PR, Brasil. E-mail: *walmes@ufpr.br*

identificação pode garantir, na maioria das vezes, o sucesso da atividade. A tendência geral do crescimento para a maioria das espécies de animais de interesse econômico é uma curva na forma de S estendido, denominada curva de crescimento, e esta representa o desenvolvimento dos animais em todas as fases de sua vida. Modelos de regressão não-linear têm se mostrado adequados para descrever estas curvas, pois apresentam parâmetros que podem ser interpretados biologicamente. Estes são caracterizados principalmente pelo peso à maturidade, que representa o peso na idade adulta, e pela velocidade de crescimento, ou taxa de maturidade, que representa uma medida de precocidade. Alguns modelos bastante utilizados na literatura são: Logístico, Gompertz, Brody, Von Bertalanffy e Richards (Freitas, 2005).

Por se tratar de modelos não-lineares, geralmente os métodos de estimação caracterizam-se como processos iterativos, e dentre estes métodos destaca-se o algoritmo de Gauss-Newton, cujo emprego é amplamente utilizado em estudos de curvas de crescimento. Um dos problemas frequentemente encontrados em tais estudos diz respeito à não convergência do algoritmo, a qual impossibilita a obtenção das estimativas dos parâmetros, e consequentemente, a identificação de indivíduos com maior eficiência de crescimento. McManus et. al. (2010) ao analisarem curvas de peso e de altura de quatro grupos genéticos de cavalos, verificaram que o modelo Richards não convergiu para nenhuma das variáveis analisadas; e que o modelo Logístico não convergiu ao ajustar curvas de peso e o de Gompertz no ajuste de altura para a maioria dos grupos considerados.

Um fator importante para a não-convergência dos modelos é a presença de oscilações na trajetória esperada da curva, caracterizadas muitas vezes por uma perda de peso abrupta dos animais decorrente da influência de efeitos ambientais como falta de nutrientes e/ou presença de enfermidades. Este fenômeno de perda de peso tem sido relatado em estudos recentes de curvas de crescimento, dentre eles, Mendes et al. (2009), envolvendo bovinos da raça Hereford criados no sul do país; Silveira (2010), envolvendo cordeiros cruzados Dorper x Rabo Largo criados na Bahia; e Sousa et al. (2010), envolvendo novilhos da raça Guzerá criados em Minas Gerais. Em todos estes trabalhos, embora se esperasse um comportamento natural crescente do peso em função da idade, alterações na forma da curva causadas por fatores externos e caracterizadas por decréscimos do peso foram observadas, e comprometeram a convergência dos modelos de regressão não-lineares utilizados.

O objetivo do presente trabalho foi propor uma metodologia de transformação de dados, via análise de regressão isotônica, para estudos de curvas de crescimento cujos dados apresentam distúrbios caracterizados por decréscimos de pesos em determinadas faixas de idades. O procedimento estatístico conhecido como regressão isotônica, cuja teoria foi sumarizada por Barlow et al. (1972), contempla a ordenação existente na variável resposta no processo de estimação. Além de investigar a eficiência da metodologia baseada em regressão isotônica em relação ao aumento da convergência e da qualidade do ajuste do modelo, objetivou-se também propor um procedimento iterativo de isotonização cujo intuito foi obter uma transformação otimizada para os dados. Todas as metodologias mencionadas foram avaliadas por meio de um estudo de simulação Monte Carlo.

2 Metodologia

2.1 Curvas de crescimento

Uma forma prática e eficiente de se analisar o crescimento de animais é por meio do estudo de suas curvas de crescimento, as quais descrevem uma relação funcional entre peso e idade (SILVA et al., 2001). Geralmente, tal relação é representada por modelos de regressão não-lineares, os quais, em sua maioria, apresentam os seguintes parâmetros (SILVA et al., 2004): peso assintótico superior, que representa o peso médio à maturidade, ou peso adulto; taxa de maturidade, que representa a velocidade de crescimento, de forma que quanto mais alto for o seu valor, mais precoce é o animal e vice-versa (Brown, Fitzhugh Jr, Cartwright, 1976); ponto de inflexão, que se refere ao momento em que o animal muda de uma fase de crescimento acelerada para uma fase de crescimento inibitória; e por último a constante integração, o qual não apresenta uma interpretação biológica direta.

Segundo Fitzhugh Jr. (1976), os seguintes requisitos devem ser atendidos para que um modelo de regressão não-linear descreva adequadamente a relação peso-idade: interpretação biológica dos parâmetros, “alta qualidade” de ajuste, e facilidade de convergência.

2.2 Regressão isotônica

Seja $X = \{x_1, x_2, ..., x_k\}$ um conjunto finito, munido de uma relação binária ($\prec$) entre os elementos de X. Tal relação binária é uma ordem simples se são satisfeitas as propriedades:

1) reflexiva: $x \prec x$ para todo $x \in X$;
2) transitiva: $x, y, z \in X,\ x \prec y \ e \ y \prec z \quad então \quad x \prec z$;
3) anti-simétrica: $x, y \in X,\ x \prec y \ e \ y \prec x \quad então \quad x = y$; e
4) todo e qualquer elemento de X é comparável: $x, y \in X$, implica que, $x \prec y \quad ou \quad y \prec x$.

Definição 1. *Uma função* $f : (X, \prec) \rightarrow R$ *é dita isotônica se*

$$x_1 \prec x_2 \Rightarrow f(x_1) \leq f(x_2)$$

Definição 2. *Uma função* $\omega : X \rightarrow R$ *positiva é dita uma função peso.*

Para o caso da ordenação simples, a regressão isotônica é definida como segue.

Definição 3. *Seja* g *uma função em* X. *Uma função* g^* *é uma regressão isotônica de* g *com peso* ω *se, e somente se,* g^* *é uma função isotônica e minimiza*

$$\sum_{x \in X}[g(x) - f(x)]^2 \omega(x) \tag{1}$$

onde f *varia entre todas as funções isotônicas em* X *. Portanto,* g^* *é uma solução de mínimos quadrados restritos para a expressão (1).*

Em várias situações práticas, devido a alguma informação “a priori”, os parâmetros de interesse apresentam algum tipo de ordenação. Assim, deve-se buscar um modelo que preserve esta característica de ordem. O método para encontrar este modelo é denominado Regressão Isotônica, sendo que o termo isotônica (ou monotônica) refere-se ao fato de que um aumento na variável independente implique em um aumento na variável resposta. Se a variável resposta decresce com o aumento da variável preditora, utiliza-se o termo antitônica. Portanto, quando é razoável supor que a variável resposta cresça com o aumento da variável independente, uma teoria adequada a ser utilizada é a da regressão isotônica.

A regressão isotônica geralmente é feita em dois tipos de conjuntos de dados; os conjuntos com uma ordenação simples das observações, e os conjuntos com uma quase ordenação das observações. No presente trabalho será considerado apenas teorias de regressão isotônica fundamentadas na ordenação simples, seguindo a notação e os conceitos apresentados no livro clássico de Barlow et al. (1972).

2.3 O algoritmo PAVA

Um algoritmo amplamente utilizado para calcular a regressão isotônica para uma ordem simples é o PAVA (*pool-adjacent-violators algorithm*). O algoritmo inicia-se com a função $g(x)$, e se $g(x)$ é isotônica, então $g(x) = g^*(x)$. Caso contrário deve existir algum índice i tal que $g(x_{i-1}) > g(x_i)$, o valor $g(x_{i-1})$ é denominado violador. Estes dois valores são substituídos pelas suas médias ponderadas, as quais são denotadas por $Av(i-1, i)$, isto é:

$$Av(i-1,i) = \frac{g(x_{i-1})\omega(x_{i-1}) + g(x_i)\omega(x_i)}{\omega(x_{i-1}) + \omega(x_i)} \tag{2}$$

Após o cálculo desta média ponderada, os pesos $\omega(x_{i-1})$ e $\omega(x_i)$ serão substituídos pela soma $\omega(x_{i-1}) + \omega(x_i)$. Se esse novo conjunto, com $p-1$ valores é isotônico, então $g^*(x_i - 1) = g^*(x_i) = (Av(i-1,i))$ e $g^*(x_j) = g(x_j)$, com $j = 1, 2, ..., k-1$. Caso contrário, existe um violador e o processo precedente é repetido até que se obtenha um conjunto com valores isotônicos. Observe que os valores de g^* são as médias ponderadas dos blocos de valores para os quais existia um violador. Desta forma, a regressão isotônica é obtida por um procedimento simples, com significado estatístico, o uso de médias locais para os valores que apresentavam violadores.

2.4 Isotonização de dados de curvas de crescimento

Considere as observações y_i referente aos pesos de um animal tomado nos tempos x_i, $i = 1,2,...,n$ e os modelos de crescimento Gompertz, Logístico e Von Bertalanffy apresentados na Tabela 1. O ajuste destes modelos ao conjunto de dados originais de peso-idade $\mathbf{y}$, por meio do método dos Quadrados Mínimos Ordinários consiste em assumir $\mathbf{y} = \mathrm{f}(\mathbf{x}, \boldsymbol{\theta}) + \boldsymbol{\varepsilon}$, em que $\mathbf{y} = [y_1, y_2, ..., y_n]$, $\mathbf{x} = [x_1, x_2, ..., x_n]$, $\boldsymbol{\varepsilon} = [\varepsilon_1, \varepsilon_2, ..., \varepsilon_n]$, $\varepsilon_i \sim N(0, \sigma^2)$ e $\boldsymbol{\theta}$ é o vetor de parâmetros do modelo.

Tabela 1 - Modelos não-lineares de crescimento adotados

Modelos adotados	Expressão Matemática
Gompertz	$y = Ae^{-be^{(-kt)}} + \varepsilon$
Logístico	$y = \dfrac{A}{1 + be^{-kt}} + \varepsilon$
Von Bertalanffy	$y = A(1 - be^{-kt})^3 + \varepsilon$

Os dados y_i, $i = 1,2,...,n$ não estão necessariamente em ordem crescente em razão de variações aleatórias. Uma transformação dos dados para se obter valores crescentes é feita via regressão isotônica, denominada de isotonização de dados. Com este procedimento os valores originais $\mathbf{y} = [y_1, y_2, ..., y_n]$ são transformados em valores $\mathbf{y}^* = [y_1^*, y_2^*, ..., y_n^*]$, crescentes em relação às datas das avaliações. Dessa forma, os modelos de crescimento serão ajustados ao conjunto de dados $\mathbf{y}^* = [y_1^*, y_2^*, ..., y_n^*]$.

Como todo processo de regressão, a curva final é muito influenciada pela ocorrência de *outliers.* A teoria da regressão isotônica é suficientemente flexível no sentido que pode se diminuir o efeito de *outliers* através do uso da regressão isotônica com pesos diferentes. Para determinar pesos adequados foi utilizado o inverso da diferença entre o valor do dado inicial y_i e o valor correspondente após a isotonização y_i^*, $\omega(x_i) = 1/|y_i - y_i^*|$, se $y_i \neq y_i^*$. Com estes pesos os dados originais são novamente isotonizados obtendo-se novos valores $\mathbf{y}^{**} = [y_1^{**}, y_2^{**}, ..., y_n^{**}]$.

A idéia para o uso de tais pesos é que se y_i^* está longe de y_i, isto é y_i é um valor discrepante em relação ao conjunto dos dados, o peso $\omega(x_i)$ será relativamente pequeno o que diminuirá sua influência na obtenção dos dados isotonizados, impedindo assim que estes *outliers* afetem por demais o ajuste crescente dos dados. No contexto prático, o que se pretende é desvalorizar uma perda excessiva de peso do animal decorrente de um fator ocasional. Tem-se, portanto, com este procedimento, que os dados $\mathbf{y} = [y_1, y_2, ..., y_n]$ são transformados em $\mathbf{y}^* = \left[y_1^*, y_2^*, ..., y_n^* \right]$ e, posteriormente, utilizando-se os pesos $\omega(x_i) = 1/|\, y_i - y_i^* \,|$, transformados em $\mathbf{y}^{**} = \left[y_1^{**}, y_2^{**}, ..., y_n^{**} \right]$.

As três situações mencionadas, dados originais, transformados via regressão isotônica com pesos iguais e via regressão isotônica com pesos diferentes podem ser visualizadas na Figura 1.

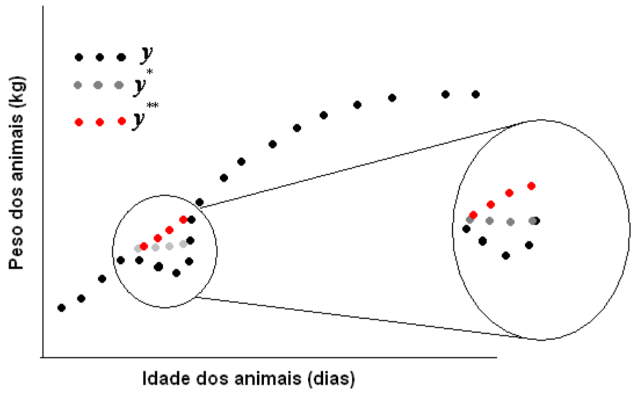

Figura 1 - Representação gráfica dos processos de isotonização efetuados.

No sentido de obter uma transformação otimizada para os dados, diminuindo a influência dos *outliers*, o processo anterior pode ser sucessivamente aplicado obtendo um processo iterativo:

Passo 1: Transformar os dados originais $\mathbf{y} = \left[y_1, y_2, ..., y_n \right]$ pela regressão isotônica com pesos iguais obtendo-se $\mathbf{y}^{*(1)} = \left[y_1^{*(1)}, y_2^{*(1)}, ..., y_n^{*(1)} \right]$.

Passo 2: Transformar os dados originais $\mathbf{y} = \left[y_1, y_2, ..., y_n \right]$ pela regressão isotônica com pesos diferentes $\omega(x_i) = 1/|\, y_i - y_i^* \,|$, se $y_i \neq y_i^*$, obtendo-se $\mathbf{y}^{*(2)} = \left[y_1^{*(2)}, y_2^{*(2)}, ..., y_n^{*(2)} \right]$.

.

.

Passo k: Transformar os dados originais $\mathbf{y} = \left[y_1, y_2, ..., y_n \right]$ pela regressão isotônica com pesos diferentes $\omega(x_i) = 1/|\, y_i - y_i^{*(k-1)} \,|$, se $y_i \neq y_i^{*(k-1)}$ obtendo-se $\mathbf{y}^{*(k)} = \left[y_1^{*(k)}, y_2^{*(k)}, ..., y_n^{*(k)} \right]$.

O processo é aplicado até que não se verifique diferenças significativas entre os valores isotonizados no ($k-1$)-ésimo passo e o k-ésimo passo. A Figura 2 tem por objetivo elucidar este processo iterativo.

.

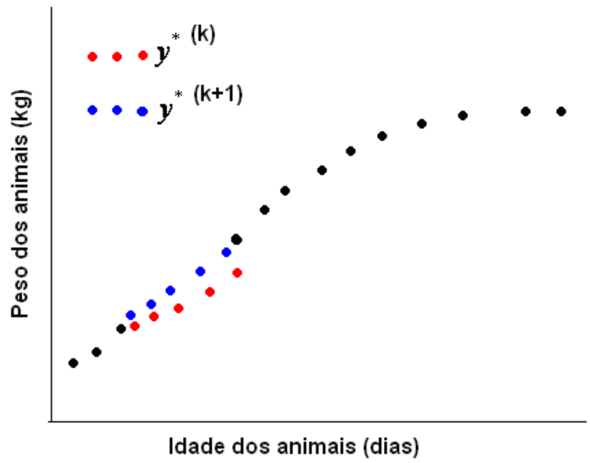

Figura 2 - Representação gráfica do processo iterativo baseado em regressão isotônica com pesos diferentes.

Em estudos de simulação de dados, é possível avaliar a eficiência deste procedimento iterativo por meio do cálculo do Erro Quadrático Médio (EQM) para cada parâmetro em cada uma das iterações, sendo o critério de eficiência a redução do EQM em relação às iterações.

É importante comentar que a isotonização de dados garante uma redução do erro de ajuste de modelos crescentes conforme teorema apresentado por Robertson, Wright e Dykstra (1988, pág. 41).

A regressão isotônica pode ser implementada em diferentes softwares estatísticos, e dentre estes o software livre R (R Development Core Team, 2009) é aquele que apresenta tal implementação de forma mais simples e compacta, e por este motivo tal software foi utilizado no presente trabalho. Neste software, o pacote requerido ao se considerar pesos iguais é o *isotone*, e a função usada foi a *isoreg*; e ao se considerar pesos diferentes, o pacote requerido foi o *cir*, e a função usada foi a *cir.pava*.

2.5 Estudo de simulação de dados

A fim de avaliar a eficiência da regressão isotônica para corrigir distúrbios em curvas de crescimento animal, foi realizado um estudo de simulação de dados via técnica Monte Carlo. O estudo de simulação em questão consistiu em gerar valores de y_i por meio da adoção de valores fixos para os parâmetros dos modelos (Gompertz, Logístico e Von Bertalanffy). Foram geradas vinte observações longitudinais, sendo x_i = 0, 30, 60,, ..., 600, pois na prática, em experimentos envolvendo curvas de crescimento animal geralmente as pesagens são obtidas mensalmente. Para cada observação y_i foi gerado um valor aleatório relacionado ao termo de erro, o qual foi especificado como sendo ε_i ~N(0,3). A variância foi fixada no valor 3 por ser compatível com os valores utilizados para os parâmetros dos modelos. Ao todo foram simuladas 1000 curvas de crescimento diferentes, sendo estas diferenças garantidas pelo número aleatório gerado para o erro. Em termos práticos, estas 1000 curvas podem corresponder a curvas de 1000 animais diferentes pertencentes a um mesmo rebanho.

Para inserir distúrbios na curva de crescimento, a fim de que a mesma não apresentasse a trajetória crescente sigmóide esperada, primeiro calculou-se a derivada discreta para cada ponto de cada curva simulada, ou seja, $y_i' = \frac{y_i - y_{i-1}}{x_i - x_{i-1}}$. Tomando-se o valor máximo da derivada discreta de cada curva, o valor de x_i correspondente a este ponto de máximo representa o tempo no qual ocorre o ponto de inflexão da curva.

Tal procedimento foi adotado visto que distúrbios de crescimento geralmente ocorrem em torno do ponto de inflexão, pois de acordo com Gottschall (1999) é onde o animal cresce com maior velocidade, e consequentemente, demanda maior qualidade de manejo (geralmente maior demanda nutricional). Assim, quando o manejo nesta fase de crescimento não é adequado, o animal pode reduzir seu peso abruptamente, pois suas necessidades diárias de alimentos não são atingidas justamente quando ele mais precisa.

De acordo com a já relatada importância do ponto de inflexão, no estudo de simulação uma função de distúrbio foi inserida tendo como referência este ponto, e para tanto foi utilizada a seguinte função paramétrica: $f(\mu,\alpha,\beta,x) = -\beta.\exp\left\{\frac{-(x-\mu)^2}{\alpha}\right\}$, em que: μ é o parâmetro que controla a posição, o qual coincide com o valor do tempo no qual ocorre o ponto de inflexão, α é o parâmetro que controla a extensão e β é o parâmetro que controla a intensidade do distúrbio. Dessa forma, a função de distúrbio foi simplesmente somada à curva de crescimento original, fato este que culminou no surgimento de um distúrbio conforme apresentado na Figura 3.

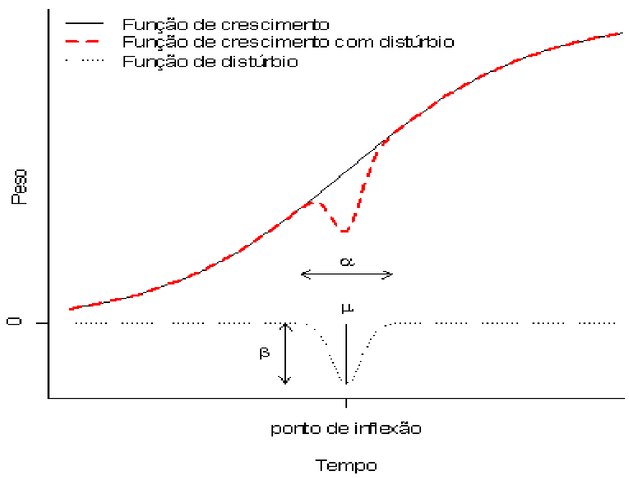

Figura 3 - Esquema ilustrativo da inserção da função de distúrbio na curva de crescimento original.

Para os cálculos foram utilizados a função *isoreg* do pacote *isotone* do software R para a regressão isotônica com pesos iguais e a função *cir.pava* do pacote *cir* do software R para a regressão isotônica com pesos diferentes.

Após a especificação das três situações (dados originais com distúrbio, dados corrigidos pela regressão isotônica com pesos iguais e dados corrigidos pela regressão isotônica com pesos diferentes), realizaram-se os ajustes de cada um dos três modelos Gompertz, Logístico e Von Bertalanffy para cada uma das situações a fim de determinar a eficiência dos procedimentos de isotonização em relação à porcentagem de convergência. Estes ajustes foram realizados por meio da função *nls* do R, que estimou os parâmetros mediante método dos mínimos quadrados ordinários via algoritmo de Gauss-Newton. Nos ajustes em questão, os valores iniciais requeridos para se iniciar o processo iterativo de estimação foram os mesmos valores dos parâmetros usados para simular os dados de peso-idade. A utilização dos próprios valores paramétricos como valores iniciais é uma forma de padronizar a execução do processo iterativo requerido pelo método de estimação. Assim, tais valores, os quais podem ser denominados de ótimos, evitam que um ou outro modelo seja prejudicado pela adoção de valores que dificultem a convergência. É claro que nas situações práticas o valor paramétrico é desconhecido. No entanto existem metodologias na literatura para obtenção de valores iniciais adequadas, dentre as quais destaca-se a apresentada por Fekedulegn et al. (1999).

Além da porcentagem de convergência, foi ainda calculado o Erro Quadrático Médio (EQM) para cada um dos parâmetros dos três modelos em todas as três situações acima mencionadas. As expressões do EQM para os parâmetros A, b e K considerando j=1, 2,...,1000 repetições da simulação, foram dadas respectivamente por:

$$\mathrm{EQM}(A) = \frac{\sum_{j=1}^{1000}\left(A - \hat{A}_j\right)^2}{1000},\quad \mathrm{EQM}(b) = \frac{\sum_{j=1}^{1000}\left(b - \hat{b}_j\right)^2}{1000}\ \ e\ \ \mathrm{EQM}(K) = \frac{\sum_{j=1}^{1000}\left(K - \hat{K}_j\right)^2}{1000},$$

em que A, b e K foram os valores paramétricos considerados na simulação das curvas.

Tendo em vista a abordagem iterativa, que consistiu em aplicar subseqüentemente várias regressões isotônicas ponderadas, é importante comentar que a eficiência de tal metodologia proposta também foi verificada pelo cálculo do EQM para cada parâmetro, porém neste contexto os EQM´s foram calculados em cada uma das iterações consideradas. No presente trabalho foi assumido o número máximo de apenas 10 iterações, devido ao custo computacional envolvido. Conforme já comentado no item citado, a eficiência deste método pode ser facilmente verificada por um decréscimo do EQM em cada iteração.

3 Resultados e discussão

A proposta de isotonização de dados referentes a estudos envolvendo curvas de crescimento animal, cujo objetivo principal foi melhorar a porcentagem de convergência (%C) e qualidade dos ajustes em situações que consideram curvas atípicas relacionadas com a perda de peso dos animais, foi avaliada por meio da simulação Monte Carlo. As três situações consideradas (dados originais com distúrbio, dados corrigidos via regressão isotônica com pesos iguais e via regressão isotônica com pesos diferentes) nesta simulação podem ser visualizadas na Figura 4, as quais retratam respectivamente, as curvas de crescimento geradas por meio dos modelos Logístico, Von Bertalanffy e Gompertz para dois indivíduos aleatoriamente amostrados da população base de 1000 indivíduos.

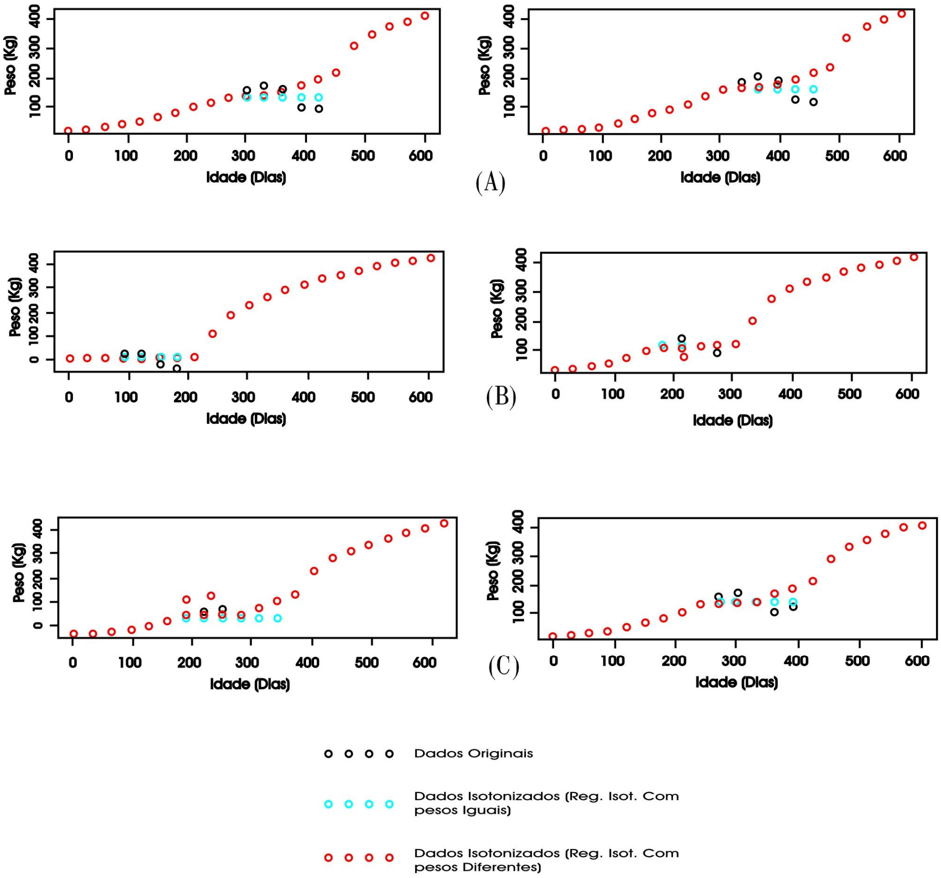

Figura 4 - Curvas de crescimento de dois animais geradas, respectivamente, via modelos Logístico (A), Von Bertalanffy (B) e Gompertz (C).

Os gráficos mostrados na Figura 4 explicitam a idéia de transformação de dados via regressão isotônica, a qual permite de certa forma, recuperar a trajetória natural da curva representativa do fenômeno em estudo, que neste caso trata-se do crescimento de animais de interesse Zootécnico. Propostas semelhantes já foram aplicadas com sucesso na literatura estatística, como a de Hussian et al. (2004), que envolveu ajuste de modelos de regressão não-linear para descrever a concentração média de nitrogênio em função do tempo; e a de Gunn & Dunson (2005), que envolveu ajuste de modelos de regressão para descrever a concentração de progesterona no decorrer de cada ciclo menstrual.

O estudo de simulação de dados evidenciou a utilidade da proposta de se utilizar a isotonização de dados em estudos de curvas de crescimento animal. As Tabelas 2, 3 e 4 mostram tais resultados, respectivamente, para os modelos Logístico, Von Bertalanffy e Gompertz. Nestas Tabelas % C significa a porcentagem de convergência.

Tabela 2 - Resultado do processo de simulação para o modelo Logístico

Situação (transformação dos dados)	%C	EQM (A)	EQM (b)	EQM (K)
Originais ($\mathbf{y}$)	50	529786605,6	426037,2492	8,170361e-06
Reg. Isot. Pesos Iguais ($\mathbf{y}^*$)	50	2192547,1	1959,7055	5,725936e-06
Reg. Isot. Pesos Diferentes ($\mathbf{y}^{**}$)	100	539594,4	265,4507	5,994183e-06

Tabela 3 - Resultado do processo de simulação para o modelo Von Bertalanffy

Situação (transformação dos dados)	%C	EQM (A)	EQM (b)	EQM (K)
Originais ($\mathbf{y}$)	80	12552,609	0,1001136	7,243446e-07
Reg. Isot. Pesos Iguais ($\mathbf{y}^*$)	100	7568,218	0,1116972	4,646744e-07
Reg. Isot. Pesos Diferentes ($\mathbf{y}^{**}$)	100	5849,840	0,1117990	3,311908e-07

Tabela 4 - Resultado do processo de simulação para o modelo Gompertz

Situação (transformação dos dados)	%C	EQM (A)	EQM (b)	EQM (K)
Originais ($\mathbf{y}$)	50	1311510,1	1,5368543	5,110558e-06
Reg. Isot. Pesos Iguais ($\mathbf{y}^*$)	80	2625149,8	0,0943173	7,153625e-06
Reg. Isot. Pesos Diferentes ($\mathbf{y}^{**}$)	100	406188,3	0,3113374	4,321749e-06

Em um âmbito geral, os resultados mostrados nas Tabelas 2, 3 e 4 sugerem que o procedimento de isotonização dos dados originais de crescimento na presença de alterações não naturais, aqui definidas como distúrbios, mostrou-se eficiente em relação ao aumento da %C para todos os modelos estudados. Tal eficiência foi evidente para a transformação baseada na regressão isotônica com pesos diferentes, a qual providenciou uma %C de 100% para todos os três modelos adotados.

Nota-se que a maior eficácia da isotonização foi constatada quando foram utilizados os modelos Logístico (Tabela 2) e Gompertz (Tabela 4), pois para estes dois modelos a %C ao se considerar os dados originais foi apenas de 50%. Sob um ponto de vista prático, ao se aplicar estes modelos na descrição de curvas de crescimento com distúrbios referentes a animais de um determinado rebanho, ter-se-ia informações de eficiência de crescimento por meio da interpretação biológica dos parâmetros apenas para a metade dos animais.

Tendo em vista os resultados da regressão isotônica com pesos iguais, nota-se que esta foi eficiente quando se considerou os modelos Von Bertalanffy (Tabela 3) e Gompertz (Tabela 4), uma vez que as %C foram 100 e 80%, respectivamente. Porém, esta forma simplificada de isotonização não conduziu a bons resultados ao se usar o modelo Logístico, sendo a %C observada apenas de 50% a qual foi a mesma apresentada pelo ajuste considerando os dados originais.

Em resumo, os resultados da Tabelas 2, 3 e 4 permitem inferir que a hipótese levantada sobre a melhora da %C via isotonização dos dados realmente foi observada, pois o processo de suavização proveniente da aplicação de regressão isotônica, simples e ponderada, possibilitou ajustar os três modelos considerados a um maior número de animais, fato este que pode melhorar consideravelmente a qualidade de pesquisas na área de Zootecnia relacionadas com a análise de curvas de crescimento.

É possível notar que o EQM para o parâmetro A (peso adulto), para todos os modelos adotados, reduziu drasticamente em função da ordem de aplicação dos três conjuntos diferentes de dados, a qual é: dados originais, dados transformados por regressão isotônica com pesos iguais e dados transformados por regressão isotônica com pesos diferentes.

Quanto ao EQM do parâmetro b (parâmetro de integração), resultados semelhantes aqueles obtidos para o parâmetro A foram observados ao se considerar o modelo Logístico (Tabela 2), porém para os demais modelos, devido a presença de diferenças constatadas a níveis de casas decimais, tal afirmação não pode ser reafirmada.

Nas tabelas sob discussão, para todos os modelos em estudo, nota-se ainda uma expressiva melhora na precisão da estimação do parâmetro K ao se usar a regressão isotônica com pesos diferentes em relação ao uso dos dados originais e aqueles provenientes da aplicação da regressão isotônica com pesos iguais. A utilização dos dados corrigidos pela regressão isotônica com pesos iguais foi mais eficiente que a utilização dos dados originais ao se considerar os modelos Logístico e Von Bertalanffy, porém o mesmo não foi verificado ao se considerar o modelo Gompertz.

O presente trabalho também objetivou propor um procedimento iterativo baseado em regressão isotônica cuja meta foi obter uma transformação eficiente para os dados, a qual está fundamentada no Teorema apresentado por Robertson, Wright e Dykstra (1988, pág. 41). As Figuras 5, 6 e 7 mostram a avaliação da eficiência do procedimento proposto mediante avaliação o EQM para os parâmetros do modelo Logístico, Von Bertalanffy e Gompertz.

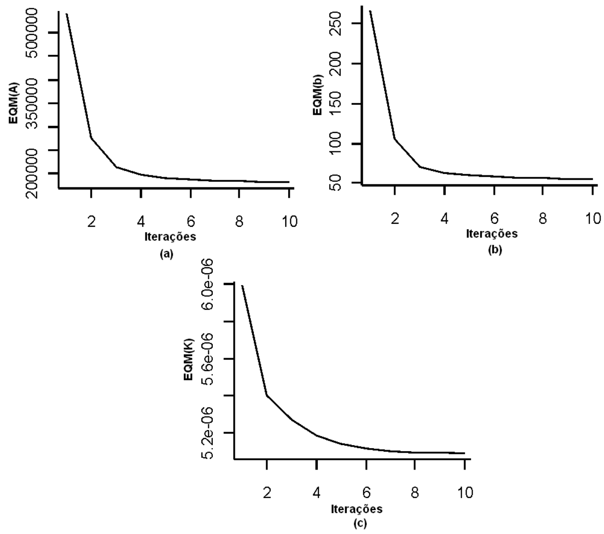

Figura 5 - Comportamento do EQM para os parâmetros A (a), b (b) e K (c) do modelo Logístico ao longo de 10 iterações consideradas pela metodologia proposta.

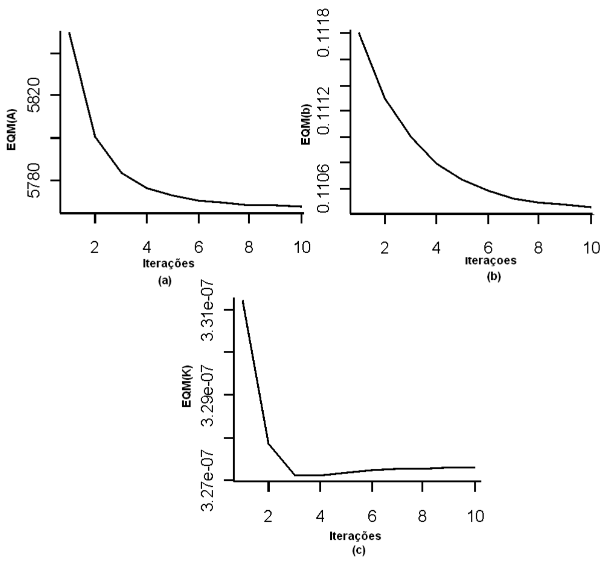

Figura 6 - Comportamento do EQM para os parâmetros A (a), b (b) e K (c) do modelo Von Bertalanffy ao longo de 10 iterações consideradas pela metodologia proposta.

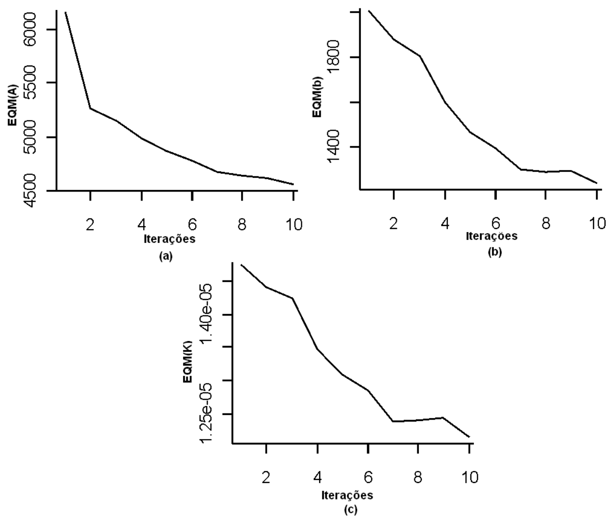

Figura 7 - Comportamento do EQM para os parâmetros A (a), b (b) e K (c) do modelo Gompertz ao longo de 10 iterações consideradas pela metodologia proposta.

De forma geral, em todas as figuras apresentadas (5, 6 e 7), observa-se um comportamento decrescente do EQM para todos os parâmetros em todos os modelos adotados em função do aumento do número de iterações. É importante realçar que para os modelos Logístico (Figura 5) e Von Bertalanffy (Figura 6) a estabilização do EQM é clara, e permite inferir que o total de 10 iterações realmente foi suficiente para se obter uma transformação ótima dos dados baseada em sucessivas aplicações da regressão isotônica com pesos diferentes. Porém, para o modelo Gompertz esta estabilização não é facilmente observada, e consequentemente ao se trabalhar com este modelo um maior número de iterações é demandado.

Além disso, o estudo de simulação empregado foi útil para verificar que a redução do EQM,ou seja, o aumento da qualidade da estimação, não é constante para diferentes modelos e nem para diferentes parâmetros destes modelos. Assim, as diferenças observadas nos gráficos das Figuras 5, 6 e 7 podem ser melhor exploradas e discutidas como se segue.

Observa-se que para os parâmetros A, Figuras 5 (a), 6 (a) e 7 (a), e b, Figuras 5 (b), 6 (b) e 7 (b), o EQM se estabiliza primeiro para o modelo Logístico, e depois para o modelo Von Bertalanffy, e por último para o modelo Gompertz, para qual inclusive a estabilização não é evidente. Para o parâmetro K o comportamento do EQM é um pouco diferente, uma vez que a estabilização é primeiramente verificada para o modelo Von Bertalanffy, seguido pelo modelo Logístico e pelo modelo Gompertz, para o qual a estabilização do EQM também não é evidenciada.

Em termos da avaliação da magnitude do EQM, menores valores foram sempre observados ao se considerar o ajuste do modelo Von Bertalanffy e para o modelo Logístico, independentemente do parâmetro avaliado.

Caso seja de interesse o estabelecimento de uma ordem referente à plausibilidade dos modelos em estudo, o modelo Von Bertalanffy conforme discutido no parágrafo anterior seria o mais plausível, sendo este seguido pelo Logístico, o qual apresenta menores valores de EQM para os parâmetros b e K em relação ao modelo Gompertz. Este último apenas apresenta menor valor de EQM para o parâmetro A, em relação ao modelo Logístico, e, além disso, ao se considerar o número máximo de 10 iterações, a estabilização do EQM não se verifica para tal modelo.

Em relação à indicação do modelo Von Bertalanffy como sendo o mais plausível tendo em vista os resultados do estudo de simulação, é relevante relatar que outros estudos de curvas de crescimento também elegeram este modelo e exaltaram suas qualidades. Silveira (2010) relatou que o modelo Von Bertalanffy foi o mais indicado para descrever curvas de crescimento de três grupos genéticos de cordeiros de corte. Guedes et al. (2004) relata que o mesmo foi o que melhor descreveu curvas de crescimento de cordeiros das raças Santa Inês e Bergamácia, e Freitas (2005) relata que o modelo em questão, juntamente com o modelo Logístico, são os mais versáteis dentre os modelos de três parâmetros para descrever curvas de crescimento de espécies domésticas.

4 Conclusão

A metodologia de transformação de dados via análise de regressão isotônica com pesos iguais e pesos diferentes, possibilitou aumentar a porcentagem de convergência e a qualidade dos ajustes dos modelos de regressão não-linear Logístico, Von Bertalanffy e Gompertz a dados de crescimento que apresentam distúrbios caracterizados por decréscimos de pesos em determinadas faixas de idades.

O método de isotonização baseado em um processo iterativo proposto no presente trabalho possibilitou a obtenção de uma transformação considerada ótima para os dados de crescimento com distúrbios.

O estudo de simulação Monte Carlo realizado possibilitou observar a eficiência das metodologias consideradas uma vez que todos os processos de isotonização adotados resultaram em maiores porcentagens de convergência e menores erros quadráticos médios (EQM) para os parâmetros dos modelos avaliados. Este estudo também possibilitou detectar diferenças na intensidade da eficiência de tais metodologias ao considerar diferentes modelos.

RODRIGUES, A.; CHAVES, L. M.; SILVA, F. F.; ZEVIANI, W. Isotonic regression applied to growth curve studies. *Rev. Bras. Biom.*, São Paulo, v.28, n.4, p.85-101, 2010.

▪ ***ABSTRACT**: Non-linear regression Models have proved to be adequate to describe growth curves of domestic animals of zootechnical interest, since they have parameters which can be interpreted from a biological point of view. These models are adjusted to weight-age data by means of iteractive algorithms such as Gauss-Newton. A problem often reported is a non convergence of this algorithm in the presence of oscillations in the expected path of the curve; it is widely characterized by an abrupt weight loss of animals due to the influence of environmental effects such as the lack of forage (nutrients) and / or the presence of diseases. Thus, it is necessary to develop estimation procedures that address this fact, and that somehow, they consider the expected* *response naturalness in the experiments. The aim of this study was to propose a methodology for processing data via isotonic regression analysis for studies of growth curves whose data have disorders characterized by decreased body weight at certain ages. Besides investigating the efficacy of the method based on isotonic regression in relation to increased convergence and the quality of the adjustment of the model. It was also intended to propose an iteractive isotonization procedure whose aim was to obtain an optimal transformation for the data. All mentioned methods were evaluated using a Monte Carlo simulation study. Through the simulation study data it was verified that the adopted isotonization methodologies resulted in higher percentages of convergence and smaller mean squared errors (MSE) for the parameters of the Logistic models, Gompertz and Von Bertalanffy.*

- ***KEYWORDS**:* Isotonic regression; nonlinear model; growth curve.

Referências

BARLOW, R. E. *Statistical inference under order restrictions*: the theory and application of isotonic regression. London: J. Wiley, 1972. 388 p.

BROWN, J. E.; FITZHUGH JR., H. A.; CARTWRIGHT, T. C. A comparison of nonlinear models for describing weight-age relationships in cattle. *J. Anim. Sci.*, Champaign, v.42, n.4, p.810-818, 1976.

FEKEDULEGN, D.; MACSIURTAIN, M.; COLBERT, J.J. Parameter estimation of nonlinear growth models in foretry. *Silvia Fennica – Research notes*. v.33, n.4, p.327-336. 1999.

FITZHUGH JR., H. A. Analysis of growth curves and strategies for altering their shapes. *J. Anim. Sci.*, Champaign, v.42, n.4, p.1036-1051, 1976.

FREITAS, A. R. de. Curvas de crescimento na produção animal. *Rev. Bras. Zootec.*, Viçosa, MG, v.34, n.3, p.786-795, 2005.

GOTTSCHALL, C. S. Impacto nutricional na produção de carne-curva de crescimento. In: LOBATO, J. F. P.; BARCELLOS, J. O. J.; KESSLER, A. M. *Produção de bovinos de corte*. Porto Alegre: EDIPUCRS, 1999. p.169-192.

GUEDES, M. H. P. et al. Estudo das curvas de crescimento de cordeiros das raças Santa Inês e Bergamácia considerando heterogeneidade de variâncias. *Ciência e Agrotecnologia*, Lavras, v. 28, n. 2, p.381-388, 2004.

GUNN, L. H.; DUNSON, D. B. A transformation approach for incorporating monotone or unimodal constraints. *Biostatistics*, New York, v.3, n.6, p.434-449, 2005.

HUSSIAN, M. et al. Monotonic regression of assessment of trends in environmental quality data. In: EUROPEAN CONGRESS ON COMPUTATIONAL METHODS IN APPLIED SCIENCES AND ENGINEERING, 1., 2004, Jyvãskylã. *Proceedings*... Jyvãskylã: ECCOMAS, 2004. p.24-28.

MCMANUS, C.; LOUVANDINI, H.; CAMPOS, V. Curvas de crescimento não-lineares para peso e altura em quatro grupos genéticos de cavalos. *Ciênc. Anim. Bras.*, Goiânia, v.11, n.1, p.80-89, 2010.

MENDES, P. N. et al. Análise da curva de crescimento difásica de fêmeas hereford por meio da função não linear de Gompertz. *Ciênc. Anim. Bras.*, Goiânia, v.10, n.2, p.454-461, 2009.

R DEVELOPMENT CORE TEAM. ***R**: a language and environment for statistical computing. Viena*: R Foundation for Statistical Computing, 2009. Disponível em: <http://www.R-project.org>. Acesso em: 25 jun. 2010.

ROBERTSON, T.; WRIGHT, F. T.; DYKSTRA, R. L. *Order restricted statistical inference*. New York: J. Wiley, 1988. 521p.

SILVEIRA, F. G. da. *Classificação multivariada de modelos de crescimento para grupos genéticos de ovinos de corte*. 2010. 59f. Dissertação (Mestrado em Estatística Aplicada e Biometria) - Universidade Federal de Viçosa, Viçosa, MG, 2010.

SILVA, F. F.; AQUINO, L. H.; OLIVEIRA, A. I. G. Influência de fatores genéticos e ambientais sobre as estimativas dos parâmetros das funções de crescimento em gado Nelore. *Ciênc. Agrotecnol.*, Lavras, v.25, n.5, p.1195-1205, 2001.

SILVA, N. A. M. da et al. Curvas de crescimento e influência de fatores não-genéticos sobre as taxas de crescimento de bovinos da raça Nelore. *Ciênc. Agrotecnol.*, Lavras, v.28, n.3, p.647-654, 2004.

SOUSA, R. C. et al. Curva de crescimento de bovinos da raça Guzerá sob prova de ganho de peso à pasto. In: SIMPÓSIO BRASILEIRO DE MELHORAMENTO ANIMAL, 8., 2010, Maringá. *Anais...* Maringá: Sbma, 2010. p.1-3.

Recebido em 11.10.2010.

Aprovado após revisão 24.01.2011.