

DIANA CAMPOS DE OLIVEIRA

FUNÇÕES SPLINES PARA ESTUDO DE CURVAS DE CRESCIMENTO EM
OVINOS CRUZADOS

Dissertação apresentada à
Universidade Federal de Viçosa,
como parte das exigências do
Programa de Pós-Graduação em
Estatística Aplicada e Biometria, para
obtenção do título de *Magister
Scientiae*.

VIÇOSA
MINAS GERAIS – BRASIL
2011

**Ficha catalográfica preparada pela Seção de Catalogação e
Classificação da Biblioteca Central da UFV**

T

O48f
2011

Oliveira, Diana Campos de, 1984-

Funções splines para estudo de curvas de crescimento em ovinos cruzados / Diana Campos de Oliveira. – Viçosa, MG, 2011.

ix, 57f. : il. (algumas col.) ; 29cm.

Inclui apêndices.

Orientador: Antonio Policarpo Souza Carneiro.

Dissertação (mestrado) - Universidade Federal de Viçosa.

Referências bibliográficas: f. 49-52.

1. Análise de regressão. 2. Ovino - Melhoramento genético - Métodos estatísticos. 3. Animais - Melhoramento genético. 4. Métodos estatísticos. 5. Modelos não-lineares (Estatística).
I. Universidade Federal de Viçosa. II. Título.

CDD 22.ed. 519.536

DIANA CAMPOS DE OLIVEIRA

FUNÇÕES SPLINES PARA ESTUDO DE CURVAS DE CRESCIMENTO EM
OVINOS CRUZADOS

Dissertação apresentada à
Universidade Federal de Viçosa,
como parte das exigências do
Programa de Pós-Graduação em
Estatística Aplicada e Biometria,
para obtenção do título de *Magister
Scientiae*.

APROVADA: 15 de fevereiro de 2011.



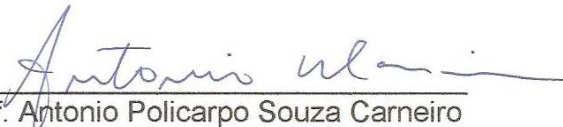
Prof. Gérson Rodrigues dos Santos



Pesq. Carlos Souza do Nascimento



Prof. Fabyano Fonseca e Silva
(Coorientador)



Prof. Antonio Policarpo Souza Carneiro
(Orientador)

AGRADECIMENTOS

Primeiramente a DEUS, por ter-me dado força para prosseguir ao longo dessa caminhada.

Aos meus pais, pelo exemplo de vida, pelo amor incondicional, por sempre estarem ao meu lado contribuindo para minhas conquistas, pelo incentivo mesmo quando eu acreditava não ser capaz.

Às minhas irmãs pela amizade, por terem estado ao meu lado sempre, apoiando-me e incentivando-me nas horas mais difíceis.

Ao Tiago pelo constante incentivo, pelo companheirismo e carinho, por sempre me apoiar em todos os momentos.

A toda minha família, que sempre torceu para que eu alcançasse meus objetivos.

Aos amigos do mestrado e graduação, pelos momentos de estudos e de diversão, pelo companheirismo e por estarem sempre presentes em todos os momentos difíceis dessa jornada, pessoas que por mais que a distância nos separe jamais os esquecerei.

Ao Professor Antonio Policarpo, pela orientação e colaboração, por sempre me receber tão disposto e prontamente sendo sua atuação de grande importância para a conclusão deste trabalho.

Aos meus co-orientadores, Professores Fabyano Fonseca e Silva, Sebastião Martins Filho, Paulo Luiz Souza Carneiro pela prontidão e conhecimentos transmitidos durante a realização desse trabalho.

À Universidade Federal de Viçosa, pela oportunidade de estudo.

A todos os professores do Departamento de Estatística da Universidade Federal de Viçosa, pelos conhecimentos transmitidos.

A Empresa Baiana de Desenvolvimento Agrícola S.A. pela concessão dos dados utilizados no trabalho.

Ao programa CAPES/REUNI, pela concessão da bolsa de estudos.

BIOGRAFIA

DIANA CAMPOS DE OLIVEIRA, filha de Antônio Klebys de Oliveira e Maria Aparecida de Oliveira, nascida em 13 de junho de 1984, na cidade de Rio Pomba, MG.

Em março de 2003, ingressou no curso de Matemática na Universidade Federal de Viçosa.

Em março de 2009, iniciou o mestrado no Programa de Pós-graduação, Estatística Aplicada e Biometria, na Universidade Federal de Viçosa, submetendo-se à defesa de dissertação em 15 de fevereiro de 2011.

SUMÁRIO

LISTA DE TABELAS.....	v
LISTA DE FIGURAS.....	vii
RESUMO.....	viii
ABSTRACT	ix
1. INTRODUÇÃO.....	1
2. REVISÃO DE LITERATURA.....	3
2.1. Ovinocultura no Brasil	3
2.2. Curvas de crescimento.....	5
2.3. Modelos não-lineares	6
2.4. Funções splines	8
2.5. Avaliadores da qualidade do ajuste.....	12
2.6. Identidade de modelos	13
3. MATERIAL E MÉTODOS.....	16
3.1 Origem dos dados	16
3.2 Obtenção dos nós da regressão spline	16
3.3 Estimação dos parâmetros da regressão spline	17
3.4 Estimação dos parâmetros dos modelos não-lineares usuais	18
3.4.1 Método dos mínimos quadrados.....	19
3.4.2 Processos iterativos.....	22
3.5 Avaliadores da qualidade de ajuste.....	24
3.5.1 Método de máxima verossimilhança	24
3.5.2 Critério de informação de Akaike (AIC)	26
3.5.3 Critério de informação bayesiano (BIC)	26
3.5.4 Coeficiente de determinação ajustado (R^2_{aj})	27
3.6. Identidade de Modelos	28
4. RESULTADOS E DISCUSSÃO	32
4.1 Estimativas dos parâmetros e dos nós das funções splines	32
4.2 Estimativas dos parâmetros da regressão não-linear	38
4.3 Avaliadores da qualidade de ajuste para regressão splines	41
4.4 Avaliadores da qualidade de ajuste para modelos não-lineares	42
4.5 Comparação entre as regressões splines e os modelos não-lineares	44
4.6 Identidade de modelos	45
5. CONCLUSÕES.....	48
6. REFERÊNCIAS BIBLIOGRÁFICAS.....	49
APÊNDICES.....	53

LISTA DE TABELAS

Tabela 1 - Modelos de regressão spline para descrever curvas de crescimento.....	18
Tabela 2 - Modelos de regressão não-linear para descrever curvas de crescimento.....	19
Tabela 3 - Estimativas dos parâmetros da regressão segmentada para o grupo DMN.....	32
Tabela 4 - Estimativas dos parâmetros da regressão segmentada para o grupo DSI.....	33
Tabela 5 - Estimativas dos parâmetros da regressão segmentada para o grupo DRL.....	33
Tabela 6 - Estimativas dos parâmetros da regressão não-linear de Richard para os três grupos genéticos.....	38
Tabela 7 - Estimativas dos parâmetros da regressão não-linear de Brody para os três grupos genéticos.....	39
Tabela 8 - Estimativas dos parâmetros da regressão não-linear de von Bertalanffy para os três grupos genéticos.....	39
Tabela 9 - Resultados do avaliador da qualidade de ajuste, critério de informação de Akaike (AIC), para as regressões splines.	41
Tabela 10 - Resultados do avaliador da qualidade de ajuste, critério de informação bayesiano (BIC), para as regressões splines.....	41
Tabela 11 - Resultados do avaliador da qualidade de ajuste, coeficiente de determinação ajustado (R_{aj}^2), para as regressões splines.....	41
Tabela 12 - Estimativas dos avaliadores da qualidade de modelos para o caso da regressão não-linear para o grupo DMN.	42
Tabela 13 - Estimativas dos avaliadores da qualidade de modelos para o caso da regressão não-linear para o grupo DRL.	43
Tabela 14 - Estimativas dos avaliadores da qualidade de modelos para o caso da regressão não-linear para o grupo DSI.	43
Tabela 15 - Estimativas dos avaliadores da qualidade de modelos para regressões splines e modelos não-lineares que apresentaram melhor	

desempenho para os grupos genéticos DMN, DRL e DSI, respectivamente.
..... 44

Tabela 16 - Testes de identidade de modelos tendo em vista os parâmetros
do modelo segmentado Linear-linear..... 45

Tabela 17 - Testes de identidade de modelos tendo em vista os parâmetros
do modelo segmentado Linear-platô..... 47

LISTA DE FIGURAS

Figura 1 - Função spline linear-linear para o grupo genético DMN.....	35
Figura 2 - Função spline linear-linear para o grupo genético DRL.....	35
Figura 3 - Função spline linear-linear para o grupo genético DSI.....	36
Figura 4 - Função spline linear-platô para o grupo genético DMN.....	36
Figura 5 - Função spline linear- platô para o grupo genético DRL.....	36
Figura 6 - Função spline linear-platô para o grupo genético DSI.....	37
Figura 7 - Função spline linear-linear-linear para o grupo DMN	37
Figura 8 - Função spline linear-linear-linear para o grupo DRL	37
Figura 9 - Função spline linear-linear-linear para o grupo DSI	38
Figura 10 - Modelo não-linear Richards para o grupo genético DMN.....	40
Figura 11 - Modelo não-linear Brody para o grupo DMN	40
Figura 12 - Modelo não-linear von Bertalanffy para o grupo DMN.....	40

RESUMO

OLIVEIRA, Diana Campos de, M.Sc., Universidade Federal de Viçosa, fevereiro de 2011. **Funções splines para estudo de curvas de crescimento em ovinos cruzados**. Orientador: Antonio Policarpo Souza Carneiro. Coorientadores: Fabyano Fonseca e Silva, Paulo Luiz Souza Carneiro e Sebastião Martins Filho.

Os objetivos deste trabalho foram utilizar a regressão spline para descrever curvas de crescimento em ovinos cruzados, comparar a qualidade de ajuste da regressão spline com modelos não-lineares usuais e aplicar teste para identidade de modelos para verificar se há diferença entre os parâmetros dos modelos de curvas de crescimento ajustadas para três grupos genéticos. Foram utilizadas, nesse trabalho, informações de pesagens de animais dos seguintes grupos genéticos de ovinos de corte: Dorper x Morada Nova (DMN), Dorper x Rabo Largo (DRL), Dorper x Santa Inês (DSI). Foram ajustadas curvas de crescimento através dos modelos não-lineares von Bertalanffy, Richards e Brody e das funções splines linear-linear, linear-platô e linear-linear-linear para cada grupo genético. As splines e os modelos não-lineares foram comparados através dos seguintes avaliadores da qualidade de ajuste: AIC (critério de informação de Akaike), BIC (critério de informação bayesiano), Coeficiente de determinação ajustado (R_{aj}^2). As splines apresentaram melhor ajuste do que os modelos não-lineares, sendo a linear-platô a mais indicada para descrever a relação de peso-idade dos ovinos em estudo. De acordo com testes de identidade de modelos, o grupo genético DSI foi o que apresentou maior precocidade e maior ganho de peso diário, sendo indicado para produção de carne.

ABSTRACT

OLIVEIRA, Diana Campos de, M.Sc., Universidade Federal de Viçosa, February, 2011. **Splines functions for the study of growth curves in crossed sheeps**. Adviser: Antonio Policarpo Souza Carneiro. Co-advisers: Fabyano Fonseca e Silva; Sebastião Martins Filho and Paulo Luiz Souza Carneiro.

This study aims to describe growth curves in crossed sheeps through spline regression, to compare the goodness of the spline regression adjustment using usual non-linear models and, then, to apply a test for identity of models in order to verify if the curves of the groups presented differences. It was used, in this work, animals weighings informations from the followings cutting sheeps genetical groups: Dorper x Morada Nova (DMN), Dorper x Rabo Largo (DRL), Dorper x Santa Inês (DSI). It was adjusted growth curves through von Bertalanffy, Richards and Brody non-linear models; and through linear-linear, linear-plateau and linear-linear-linear spline functions for each genetical group. The spline functions and the non-linear models were compared by the following quality of fit evaluators: AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion), Adjusted Determination Coefficient (R^2_{aj}). The spline functions presented better adjustment than the non-linear models, and the linear-plateau was the best one to describe the weight-age relation of the sheeps studied. According to tests of models identity, the genetical group DSI presented greater precocity and higher daily weight gain; then, this group was the most suitable for meat production.

1. INTRODUÇÃO

A ovinocultura de corte no Brasil, atualmente, apresenta grande mudança. O mercado de carne ovina está em grande expansão devido ao aumento na quantidade de consumidores e produtores. Os preços hoje são mais representativos do que o preço pago pela carne bovina nas mesmas condições. Apesar disto, boa parte da carne ovina consumida no Brasil é proveniente de outros países como Argentina, Uruguai e Nova Zelândia. Isso mostra uma possibilidade enorme de mercado a ser conquistado pelos produtores brasileiros. Principalmente, porque o Brasil tem grande potencial para produzir carne de ótima qualidade e baixo custo (FRANÇOIS, 2009). Sendo assim, a produção brasileira poderá atender à demanda interna e no futuro o mercado internacional.

Uma forma, consistente e prática, de analisar a eficiência produtiva de animais é por meio do estudo de curvas de crescimento as quais relacionam o peso do animal com a sua idade. Os modelos não-lineares usados para descrever tais curvas sintetizam um grande número de medidas em apenas alguns parâmetros (SILVA et al., 2001).

Os criadores de animais de corte estão cada vez mais preocupados em obter animais de crescimento rápido e que atinjam condição corporal para abate mais cedo, visando reduzir os custos de produção e aumentar a rentabilidade.

No melhoramento animal, o objetivo é avaliar os fatores genéticos e de ambiente que influenciam a forma da curva de crescimento e, desse modo, alterá-la por meio de seleção, intensificando as características economicamente importantes. Por exemplo, identificando animais com maior velocidade de crescimento, sem alterar o peso adulto, em vez de selecionar animais apenas para peso.

Modelos não-lineares têm sido muito usados para descrever curvas de crescimento, mas regressões segmentadas conectadas por nós (pontos de junção), as quais são denominadas funções splines, também podem ser usadas para estimar tais curvas. Essas funções apresentam facilidade de convergência, são fáceis de manipular e avaliar, são úteis quando um

polinômio de grau menor não se ajusta bem aos dados e o fenômeno em estudo apresenta grandes alterações.

De acordo com Paulson (2007), o uso de funções splines é interessante quando um modelo polinomial de pequeno grau não ajusta precisamente aos dados, e o pesquisador não quer usar uma função polinomial complexa. Em tais casos, as funções splines apresentam-se como excelente opção.

A identidade de modelos possibilita verificar se há diferença estatística no peso adulto, ganho de peso diário, e a idade em que o animal atingiu a maturidade, com isso pode-se interferir sobre a eficiência do crescimento e da produção de carne.

Segundo Regazzi e Silva (2004), em análise de regressão, com muita frequência, deseja-se saber se um conjunto de g equações ajustadas são idênticas, ou seja, se o fenômeno em estudo pode ser representado por uma única equação.

Na literatura não há aplicação de identidade de modelos para comparar funções splines.

Os objetivos deste trabalho foram avaliar a eficiência da função spline, em relação aos modelos não-lineares usuais, para descrever curvas de crescimento para três grupos genéticos de ovinos cruzados e testar a identidade de modelos das funções splines com o intuito de identificar o grupo genético com maior eficiência de crescimento.

2. REVISÃO DE LITERATURA

2.1. Ovinocultura no Brasil

Os ovinos, assim como grande maioria dos animais domésticos, foram introduzidos no Brasil principalmente pelos portugueses. Foi uma das primeiras espécies de animais domesticadas pelo homem. A sua criação possibilitava obter alimento, principalmente pelo consumo da carne e do leite, e proteção, pelo uso da lã.

Os animais eram criados somente para subsistência, principalmente na região Nordeste, mas hoje, havendo uma difusão da espécie, o Brasil conta com 16.628.571 cabeças de ovinos (IBGE, 2008). Segundo Viana (2008), tal fato se deve ao grande poder de adaptação dos ovinos a diferentes climas, relevos e vegetações.

Ainda de acordo com Viana (2008), os grandes rebanhos estão concentrados nos países localizados na Ásia, África e Oceania. No Brasil pode-se encontrar o maior rebanho na região Nordeste, seguida pela região Sul e Sudeste.

Existe clara vantagem na criação de ovinos em relação à pecuária de corte, pois fornece um maior retorno de quilos de carne por hectare em um menor espaço de tempo. Em um hectare, pode-se produzir, por ano, em torno de 500 kg de peso vivo de ovinos. Porém, se a mesma área receber bovinos, a produção será de apenas 150 kg de peso vivo (GONÇALVES, 1996). Além disso, os preços da carne ovina hoje são mais representativos do que o preço pago pela carne bovina nas mesmas condições.

As tendências para o mercado de ovino são promissoras, a demanda de carne vem sendo impulsionada pelo crescimento demográfico, pela urbanização, e pelas variações das preferências e dos hábitos alimentares dos consumidores.

Após a divulgação das qualidades típicas da carne ovina, pelo seu sabor e valor nutricional, houve um aumento considerável no consumo deste produto, o que tem ocasionado um incremento considerável em sua demanda. Apesar disso, segundo Oliveira (2010) o consumo brasileiro é 700

gramas per capita anual, o que é considerado pequeno perante a Argentina que apresenta 1,4 kg, Europa com 27 kg e a Nova Zelândia e Austrália com 42,2 kg e 20,2 kg por habitante por ano, respectivamente, sendo o Sul e o Sudeste as grandes regiões consumidoras do Brasil.

De acordo com Viana (2008), a introdução de raças especializadas, melhoramento genético e técnicas de manejo propiciaram a elevação da produtividade. Ainda assim o rebanho ovino é insuficiente para suprir o mercado interno brasileiro, desta forma, o país continua importando carne, principalmente do Uruguai, Argentina e Nova Zelândia.

Isso demonstra que existe um mercado interno ávido por consumir a carne destes animais. Também demonstra que existem oportunidades reais para os produtores brasileiros desenvolverem a cadeia produtiva de ovinos e assim ocupar de forma mais racional as imensas áreas de pasto disponíveis no país (SORIO, 2010). Entretanto, segundo Malhado et al. (2009), um dos grandes problemas para a prática desta atividade é que grande parte dos produtores não está consciente das vantagens de produzir carne de boa qualidade, e abastecem o mercado com carcaças sem qualidade, derivadas do abate de animais com idade avançada.

Uma das raças de ovinos com grande utilização para produção de carne com qualidade é a Dorper, originária da África do Sul, que surgiu do cruzamento da ovelha Blackhead Persian com o Dorset Horn (SOUZA e LEITE, 2000). Essa raça necessita de poucas exigências nutricionais e grande fertilidade, além de apresentar taxa crescimento elevada e grande adaptação às diversas condições climáticas.

A raça Morada Nova compõe uma das principais raças de ovinos deslanados explorada para carne, representa fonte de proteína na alimentação da população rural. Já a raça Rabo Largo apresenta um efetivo reduzido, é considerada umas das raças ovinas mais rústicas do Nordeste (MALHADO et al., 2009).

2.2. Curvas de crescimento

Os principais objetivos no ajuste de curvas de crescimento são descritivos, a informação contida numa seqüência de pontos de peso-idade é reduzida em poucos parâmetros que são utilizados para predizer taxas de crescimento, necessidades alimentares, peso à maturidade e graus de maturidade. Tais curvas permitem conhecer o processo de ganho de massa corporal do animal, informações que irão contribuir com a produção de carne.

As curvas de crescimento são importantes para pesquisas e recomendações sobre eficiência de produção. Pois refletem a relação entre a idade do animal e o seu impulso de crescimento e maturidade, contribuindo com a redução dos custos de produção e gerando maior rentabilidade ao produtor.

Este fato associado às características produtivas dos animais pode ser uma alternativa para programas de seleção, visando à precocidade com maior peso e melhor qualidade de carcaça (SOUZA e BIANCHINI SOBRINHO, 1994).

O crescimento dos animais tem uma grande relação com a quantidade e a qualidade da carne. Dessa forma é fundamental o conhecimento do processo de ganho de massa corporal do animal, pois esse conhecimento possibilita que se faça um controle da produção de carne. Este ganho pode ser influenciado pela alimentação, por condições climáticas, pelo estado sanitário e pelas características genéticas associadas aos animais (GOTTSCHALL, 2010).

Alguns modelos de regressão não-linear, tais como Gompertz, Richards, Brody, von Bertalanffy têm se mostrado adequados para descrever as curvas de crescimento. Pois estes modelos apresentam parâmetros que podem ser interpretados biologicamente (ELIAS, 1998).

Com relação a estudos de curvas de crescimento de ovinos, Lobo et al. (2006) ajustaram os modelos von Bertalanffy, Brody, Gompertz, Richards e Logístico a dados de peso-idade de ovinos da raça Santa Inês e o modelo Richards foi que melhor se ajustou aos dados. Entretanto, Guedes et al.

(2004) ajustaram estes mesmos modelos aos dados de peso-idade de cordeiros da raça Bergamácia e Santa Inês e concluíram que os modelos von Bertalanffy e Gompertz proporcionaram o melhor ajuste.

Silveira (2010) concluiu que os modelos Richards e von Bertalanffy foram, respectivamente, os que apresentaram os melhores ajustes para os conjuntos de dados médios e individuais de ovinos dos mesmos grupos genéticos utilizados nesse trabalho.

Além dos modelos não-lineares, as funções splines também podem ser usadas para relacionar peso-idade de animais, através de tais funções é possível ajustar modelos diferentes para diferentes intervalos de idade dos animais.

2.3. Modelos não-lineares

Regressão é uma das técnicas estatísticas mais utilizadas para análise de dados, essa técnica permite modelar e investigar a relação entre as variáveis dependentes e independentes.

Existem muitas situações nas quais não é desejável, ou mesmo possível, descrever um fenômeno através de um modelo de regressão linear. Este fato impulsionou estudos de novas técnicas estatísticas surgindo, então, uma nova classe de modelos, os modelos de regressão não-linear, onde não são satisfeitas as condições de linearidade nos parâmetros (REGAZZI, 2003).

Ainda de acordo com Regazzi (2003), nos seres vivos o processo de crescimento passa por uma fase de rápido desenvolvimento na idade inicial e com o passar do tempo a taxa de crescimento vai decrescendo até se estabilizar na idade adulta. Desta forma funções curvilíneas assintóticas podem representar bem a relação entre as variáveis. Nesse caso, modelos não-lineares podem explicar melhor o processo de crescimento.

Um modelo é dito não-linear quando ele não é linear em relação aos parâmetros, ou seja, se pelo menos uma das derivadas parciais em relação

ao parâmetro é função de parâmetros desconhecidos, e quando não podem ser linearizado por meio de transformações.

Dessa forma, para verificar se o modelo $y_i = f(x_i, \theta) + \varepsilon_i$ $i = 1, \dots, n$ é não-linear para o parâmetro θ , é preciso verificar se a derivada $\frac{\partial f}{\partial \theta}$ depende de θ .

A relação variável resposta/variável independente, nos modelos não-lineares pode ser representada pela equação:

$$y_i = f(x_i, \theta) + \varepsilon_i, i = 1, \dots, n$$

em que:

y_i representa a observação da variável dependente;

$f(x_i, \theta)$ é a função de regressão ou função resposta;

x_i representa a observação da variável independente;

$\theta = [\theta_1, \theta_2, \dots, \theta_p]^t$ é um vetor de parâmetros p dimensional desconhecidos;

ε_i representa o efeito do erro aleatório não observável suposto NIID com média zero e variância desconhecida σ^2 .

Uma vantagem dos modelos não-lineares, em relação aos lineares, é que fornecem um bom ajuste com menos parâmetros que podem ser interpretados biologicamente, isto simplifica e facilita a interpretação em muitas situações, além de fornecer um maior conhecimento sobre o fenômeno em estudo.

Considere como exemplo o modelo não-linear de Richards:

$$y_i = \frac{\beta_1}{(1 + e^{(\beta_2 - \beta_3 x_i)})^{\frac{1}{\beta_4}}} + \varepsilon_i$$

em que:

β_1 representa o peso adulto, ou peso assintótico do animal;

β_2 não apresenta interpretação, sendo este uma constante de integração;

β_3 representa a taxa de maturidade, é um indicador da velocidade com que o animal se aproxima do peso adulto ou velocidade de crescimento (valores altos indicam maturidade precoce e valores baixos, maturidade tardia);

β_4 é denominado de parâmetro de inflexão, refere-se ao ponto em que o animal passa de uma fase de crescimento acelerada para uma fase de crescimento inibitória (DUARTE, 1975).

Para que o modelo não-linear descreva bem o processo de crescimento é importante que o ajuste apresente pequenos desvios, e que os parâmetros estejam de acordo com a realidade em estudo. Além disso, é preciso que o modelo tenha facilidade de convergência, uma vez que para estimação dos parâmetros são utilizados métodos iterativos, sendo um dos mais utilizados o de Gauss-Newton.

Tal método, segundo Regazzi (2003), pode convergir muito lentamente em alguns casos, exigindo muitas iterações ou até mesmo não convergir.

2.4. Funções splines

Outra alternativa para descrever a relação entre peso-idade, através de curvas de crescimento, é definir dois ou mais polinômios cada um para um particular intervalo de idade, ou seja, ajustar um modelo local com ajustes locais para cada segmento. Este modelo é chamado função spline ou regressão segmentada.

Segundo Hernandez (2010), as funções splines começaram a receber atenção em 1950, no controle de qualidade, onde o objetivo era determinar se uma mudança no meio havia ocorrido e onde ocorreu. Ao longo dos anos, o campo tornou-se muito mais diversificado acarretando expansão na utilização das funções splines.

A origem do nome spline vem de uma régua elástica, usada em desenhos de engenharia, que pode ser curvada de forma a passar por um dado conjunto de pontos (x_i, y_i) . Os nós que eram fixados nas áreas de interesse causando a deformação da estrutura. Apenas na década de 60 sua formulação matemática foi desenvolvida, apesar de ser usada desde o século passado.

Segundo Keele (2008), splines são funções de regressão segmentadas em que impõem restrições para que os pontos de cada segmento se unam nos pontos chamados nós.

Equação geral da spline:

$$y = \sum_{j=0}^n \beta_{0j} x^j + \sum_{i=1}^c \beta_i (x - x_0)^n,$$

em que:

- n é o grau da função spline;
- C é o número de nós;
- x_0 é o nó;
- $(x - x_0) = \begin{cases} x - x_0, & \text{se } x > x_0 \\ 0, & \text{se } x \leq x_0 \end{cases}$

A função acima definida é contínua, já que se trata de um polinômio que é contínuo em todo seu domínio e no ponto de junção temos:

$$f(x_0) = \sum_{j=0}^n \beta_{0j} x_0^j + \sum_{i=1}^c \beta_i (x_0 - x_0) = \sum_{j=0}^n \beta_{0j} x_0^j, \text{ para } x > x_0$$

$$f(x_0) = \sum_{j=0}^n \beta_{0j} x_0^j + \sum_{i=1}^c \beta_i (0) = \sum_{j=0}^n \beta_{0j} x_0^j, \text{ para } x \leq x_0$$

Em estudo de fenômenos que apresentam constantes mudanças é importante ter modelos estatísticos que modelem tais alterações. De acordo com Hernandez (2010), regressão segmentada é uma ferramenta estatística para a modelagem de fenômenos em transição e mudança. O ponto na variável explicativa em que altera a forma funcional, denominado de nó, geralmente é de algum interesse, isso porque representa uma mudança no padrão dos dados. A região delimitada pelo nó é muitas vezes referida como fases ou etapas.

Nas funções splines, em estudo de curvas de crescimento, o nó pode representar a idade em que foi atingida a maturidade ou ocorreu o desmame. Ou seja, quando houve alterações no comportamento do crescimento do animal.

A teoria e as aplicações dessa técnica são relativamente recentes, mas vem ocorrendo um rápido crescimento na sua utilização. Isso devido a sua grande utilidade em situações aplicadas e desejáveis propriedades como facilidade em avaliar e manipular, o que a torna aplicada em diversas áreas do conhecimento (SCHENKEL, 1989).

No melhoramento genético animal, as funções splines, têm sido utilizadas principalmente na análise de dados longitudinais, tais como no ajuste de curvas de lactação e crescimento, onde tem grande utilidade devido à facilidade nas interpretações dos parâmetros. Além de serem adequadas quando se trabalha com dados biológicos que são de natureza oscilatória.

Como exemplo, considere a função spline linear-linear para descrever curvas de crescimento:

$$y = a + bx, \text{ para } x \leq x_0$$
$$y = a + bx + c(x - x_0), \text{ para } x > x_0$$

Para o primeiro segmento ($x \leq x_0$), o parâmetro a que é o intercepto, ou seja, peso ao nascimento, b é a inclinação (ganho de peso diário), para o segundo intervalo ($x > x_0$), c representa a mudança na inclinação do primeiro ao segundo intervalo, se c for positivo indica aumento no ganho de peso diário, caso contrário houve um decréscimo. Sendo assim, $b+c$ é a inclinação após a mudança de intervalo, representando o novo ganho de peso diário, e x_0 é ponto ou idade de mudança na taxa de crescimento, o nó.

De acordo com Varga (2009), as grandes vantagens no uso de uma função polinomial são tornar possível a obtenção de uma curva que passa por todos os pontos conhecidos, facilidade na obtenção das derivadas e integrais da curva gerada, o que torna os polinômios viáveis na interpolação. Porém, a natureza oscilatória dos polinômios de alto grau restringe o seu uso. Tal fato não ocorre nas splines onde a aproximação polinomial de menor grau é feita para cada intervalo do domínio, conseqüentemente há pouca flutuação na curva, além de permitir curvaturas bem diferentes em cada região.

Polinômios de grau n descrevem bem a relação entre as variáveis, mas no caso de n grande há um aumento nos números de parâmetros dificultando a explicação biológica, com isso é desejável que n seja pequeno.

Segundo Schenkel (1989), uma grande qualidade que as splines apresentam é a quebra da multicolinearidade pela inclusão de zeros na coluna das variáveis explicativas, aumentando a precisão dos valores estimados, o comportamento da função em uma região pode apresentar características distintas de outras regiões, são de fácil estimação e simples interpretação biológica.

De acordo com Paulson (2007), o uso de funções splines é interessante quando um modelo polinomial de pequeno grau não se ajusta precisamente aos dados, e o pesquisador não quer usar uma função polinomial complexa. Em tais casos, as funções splines apresentam-se como excelente opção.

A maior dificuldade para o uso de funções splines é a identificação dos nós, estes podem ser estimados juntamente com os parâmetros do modelo, caracterizando um modelo não-linear ou determinado pelo pesquisador de acordo com o conhecimento prévio do fenômeno em estudo e também através da análise de diagrama de dispersão (SCHENKEL 1989).

As splines, além das curvas de crescimento, têm aplicações em ajuste de curvas de lactação, modelagem de efeitos não-genéticos, avaliações genéticas, avaliação de tendência genética e podem ser usadas em outras áreas tais como engenharia, ciências agrárias, computação e arquitetura.

Ainda em relação à utilização de splines, Robbins et al. (2005) compararam avaliações genéticas de dados de crescimento em bovinos da raça Gelbvieh utilizando funções splines lineares com três nós com o intuito de determinar a praticidade de implementação de modelos longitudinais. Os autores concluíram que modelos longitudinais podem ser eficientemente implementados por funções splines.

Huisman et al. (2002) realizaram estudos com polinômio ortogonal de Legendre cúbico e spline cúbica a fim de avaliar qual modelo melhor se ajusta aos dados de peso-idade de suínos. No entanto, os autores não chegaram à conclusão de qual foi o melhor, pois para o polinômio ortogonal

de Legendre o critério de informação Bayesiano (SCHWARZ 1978) foi superior, enquanto para a função splines cúbica o AIC (critério de informação de Akaike) foi superior.

Ainda são poucos os estudos de curvas de crescimento utilizando funções splines.

2.5. Avaliadores da qualidade do ajuste

Segundo Regazzi (2003) ao determinar a relação entre as variáveis, por meio de estudos de regressão, há o problema da definição da forma matemática que será ajustada para representar o fenômeno em estudo.

Freqüentemente, ajusta-se mais de um modelo e com base nos resultados dos avaliadores da qualidade do ajuste conclui-se qual melhor modelo.

O que geralmente encontra-se na literatura, envolvendo comparações entre modelos, é o uso apenas do coeficiente de determinação ajustado. Mas de acordo com Regazzi (2003), esse avaliador de ajuste pode não ser um bom método de comparação, uma vez que suas estimativas, muitas vezes, são muito próximas e com valor alto, perto de 100%, ainda que haja grande discrepância entre os valores estimados e observados.

Dessa forma, o que se deve fazer é utilizar também outros avaliadores da qualidade do ajuste como o critério de informação bayesiano, BIC, (SCHWARZ, 1978); critério de informação de Akaike, AIC, (AKAIKE, 1974), erro de predição médio (EPM), percentual de convergência, entre outros.

Dessa forma, segundo Silveira (2010), a seleção do modelo mais adequado, com a utilização de vários avaliadores da qualidade do ajuste, fica mais precisa, uma vez que alguns dos avaliadores levam em consideração fatores como o grau de parametrização dos modelos comparados (leva em conta o princípio da parcimônia), e outros avaliadores consideram a análise da independência residual e outros.

De acordo com Pachepsky et al. (1996), o melhor modelo é aquele que possuir pequeno número de parâmetros nas suas estimativas

apresentando uma forma mais simplificada, além de estar em acordo com leis físicas e biológicas, apresentar os menores desvios entre valores observados e preditos, e fornecer variância dos estimadores mínima.

2.6. Identidade de modelos

A verificação da identidade de modelos de regressão é uma técnica adotada para verificar se existe igualdade entre os parâmetros de duas ou mais equações estimadas.

Segundo Regazzi (2003), quando são realizados estudos com diferentes fatores ou tratamentos para verificar a relação entre as variáveis dependentes e independentes e obtêm tantas equações quanto os números de tratamentos distintos, dessa forma encontra-se o problema de verificar se certo conjunto de curvas é idêntico.

Segundo Regazzi e Silva (2004), em análise de regressão, com muita frequência, deseja-se saber se um conjunto de g equações ajustadas são idênticas, ou seja, se o fenômeno em estudo pode ser representado por uma única equação.

Na literatura são propostos vários métodos para a avaliação de identidade de modelos. Dentre estes, destacam-se o teste de razão de verossimilhança, análise de agrupamento, as variáveis indicadoras e análise de variância precedida por testes de comparações múltiplas.

Na literatura são poucos os trabalhos que utilizam identidade de modelos para comparar curvas de crescimento.

Silveira (2010) utilizou teste de identidade de modelos para verificar se as equações de três grupos genéticos de ovinos eram iguais. A autora concluiu que havia um único modelo para descrever o crescimento dos três grupos genéticos estudados.

Para o teste da razão de verossimilhança é usado o modelo que terá os parâmetros testados acrescido de uma variável indicadora (*dummy*) que representa cada um dos três grupos genéticos.

Considere g grupos ou tratamentos, o seguinte modelo denominado completo pode ser definido:

$$y_{ij} = \sum_{j=1}^g D_j [a_j + b_j x_i + c_j (x_i - x_{0j}) + d_j (x_i - x_{1j})] + e_i,$$

em que:

$a_j, b_j, c_j, d_j, x_{0j}, x_{1j}$ são os parâmetros do modelo para cada um dos g grupos genéticos, sendo $j=1, 2, \dots, g$;

e_i é erro aleatório, $e_i \sim N(0, \sigma^2)$;

D_j é uma variável indicadora, sendo:

$$D_j = \begin{cases} 1, & \text{se a observação } y_{ij} \text{ pertence ao grupo } j; \\ 0, & \text{caso contrário.} \end{cases}$$

Para ajuste do modelo completo apresentado anteriormente, tal método permite testar as seguintes hipóteses, separadamente:

$H_0^{(1)}: a_1 = a_2 = a_3$ vs $H_a^{(1)}$: nem todos a_j são iguais;

$H_0^{(2)}: b_1 = b_2 = b_3$ vs $H_a^{(2)}$: nem todos b_j são iguais;

$H_0^{(3)}: c_1 = c_2 = c_3$ vs $H_a^{(3)}$: nem todos c_j são iguais;

$H_0^{(4)}: d_1 = d_2 = d_3$ vs $H_a^{(4)}$: nem todos d_j são iguais;

$H_0^{(5)}: x_{01} = x_{02} = x_{03}$ vs $H_a^{(5)}$: nem todos x_{0j} são iguais.

Para ilustrar o teste da hipótese $H_0^{(1)}$, considere o seguinte modelo reduzido:

$$y_{ij} = \sum_{j=1}^g D_j [a + b_j x_i + c_j (x_i - x_{0j}) + d_j (x_i - x_{1j})] + e_i,$$

Se H_0 for rejeitada existem diferenças entre os grupos quanto ao parâmetro a , sendo assim são feitas comparações duas a duas para verificar tal diferença, e nesse caso as hipóteses são dadas por:

$H_0^{(7)}: a_1 = a_2$ vs $H_a^{(7)}: a_{11} \neq a_{12}$;

$H_0^{(8)}: a_1 = a_3$ vs $H_a^{(8)}: a_{11} \neq a_{13}$;

$H_0^{(9)}: a_2 = a_3$ vs $H_a^{(9)}: a_{12} \neq a_{13}$;

3. MATERIAL E MÉTODOS

3.1 Origem dos dados

Os dados utilizados neste estudo foram cedidos pela Universidade Estadual do Sudoeste da Bahia e são oriundos de um experimento realizado na Estação Experimental de Jaguaquara, pertencente à Empresa Baiana de Desenvolvimento Agrícola S.A., localizada no Município de Jaguaquara, BA, microrregião administrativa de Jequié, BA.

Foram avaliados os pesos de ovinos cruzados, compreendendo 156 animais (machos e fêmeas) dos cruzamentos Dorper x Morada Nova (DMN); Dorper x Rabo Largo (DRL) e Dorper x Santa Inês (DSI), no período de 2003 a 2005.

Foram utilizados 36 animais DMN (26 machos e 10 fêmeas), 61 DRL (36 machos e 25 fêmeas) e 59 DSI (31 machos e 28 fêmeas). Após o desmame, foram mantidos em sistema de produção semi-intensivo, durante o dia com pastejo e sal mineral, sendo recolhidos ao final da tarde.

Foram realizadas pesagens do nascimento até os 210 dias de idade em intervalos de aproximadamente 15 dias. Para ajuste dos modelos, considerou a média dos pesos por idade em função da grande irregularidade nos intervalos de pesagens.

3.2 Obtenção dos nós da regressão spline

Os nós das regressões splines podem ser estimados ou determinados por meio de diagrama de dispersão ou conhecimento da área. Tem-se um modelo não-linear quando os nós são estimados e um modelo linear se os nós forem determinados previamente pelo pesquisador.

Como exemplo para o caso linear-linear, tem-se:

$$y = a + bx, \text{ se } x \leq x_0;$$
$$y = a + bx + c(x - x_0), \text{ se } x > x_0$$

No caso em que o nó é estimado ele torna-se um parâmetro, desta forma, ao derivar em relação a x_0 a derivada fica em função de β_2 , caracterizando um modelo não-linear.

Quando o nó é pré-definido ele torna-se uma constante e as derivadas parciais em relação aos parâmetros ($a, b, c...$) não é função de nenhum parâmetro desconhecido, tem-se assim um modelo linear.

Para ajuste de regressão spline neste trabalho os nós foram ajustados junto aos demais parâmetros do modelo. Por meio do método dos mínimos quadrados cujas soluções foram obtidas por meio do processo iterativo de Gauss-Newton, através do PROC NLIN do software SAS[®] (SAS, 2003). Foi considerado um nó no caso das splines linear-linear, linear-platô, e dois nós para as splines linear-linear-linear.

3.3 Estimação dos parâmetros da regressão spline

Os nós foram estimados caracterizando uma regressão segmentada não-linear. Assim, para ajustar os parâmetros foi utilizado o método dos mínimos quadrados para modelos não-lineares.

Há muitos procedimentos numéricos para resolver tal problema, um muito utilizado é o método iterativo de Gauss-Newton cujas soluções são obtidas por meio de algoritmos computacionais que se baseiam numa aproximação por série de Taylor de primeira ordem para produzir uma linearização da função não-linear.

Embora amplamente usado, o método de Gauss-Newton pode apresentar problemas como exigir muitas iterações ao convergir, tornando lento o processo ou até mesmo não convergir, além disso, pode aumentar a soma de quadrado residual (REGAZZI, 2003). Uma alternativa para esse problema é utilizar o método de Gauss-Newton modificado disponível no PROC NLIN do software SAS[®] (SAS, 2003).

Grande parte de trabalhos com modelos de regressão segmentada são feitos no contexto dos mínimos quadrados, no entanto, segundo Hernandez (2010) quando os dados são normalmente distribuídos as estimativas através dos mínimos quadrados são equivalentes com as estimativas de máxima verossimilhança.

Neste trabalho, foram estimados os parâmetros para as funções splines linear-linear, linear-platô e linear-linear-linear (Tabela 1), essa escolha se deve ao comportamento dos dados analisados pelo diagrama de dispersão.

Tabela 1 - Modelos de regressão spline para descrever curvas de crescimento de ovinos cruzados

Splines	Equações
Linear-platô	$y_i = a + bx_i$, se $x < x_0$
	$y_i = a + bx_0$, se $x \geq x_0$
Linear-linear	$y_i = a + bx_i$, se $x < x_0$
	$y_i = a + bx_i + c(x-x_0)$, se $x \geq x_0$
Linear-linear-linear	$y_i = a + bx_i$, se $x < x_0$
	$y_i = a + bx_i + c(x-x_0)$, se $x_0 < x \leq x_1$
	$y_i = a + bx_i + c(x-x_0) + d(x-x_1)$, se $x > x_1$

3.4 Estimação dos parâmetros dos modelos não-lineares usuais

Também foram ajustados os modelos usuais de regressão não-linear Richards, Brody e von Bertalanffy para os três grupos genéticos de ovinos. A escolha destes modelos se deve ao fato de que em trabalhos anteriores, realizados com os mesmos grupos genéticos, esses modelos foram os que apresentaram melhor desempenho (Tabela 2).

Tabela 2 - Modelos de regressão não-linear para descrever curvas de crescimento

Modelo de Curva de Crescimento	Referência
$y_i = \frac{\beta_1}{\left(1 + e^{(\beta_2 - \beta_3 x_i)}\right)^{\beta_4}} + e_i$	Richards (1959)
$y_i = \beta_1 \left(1 - \beta_2 e^{-\beta_3 x_i}\right) + e_i$	Brody (1945)
$y_i = \beta_1 \left(1 - \beta_2 e^{-\beta_3 x_i}\right)^3 + e_i$	von Bertalanffy (1957)

Para esses modelos (Tabela 2) tem-se que y_i representa a observação da variável dependente (peso), enquanto x_i representa a observação da variável independente (idade), β_1 o peso adulto ou peso assintótico do animal, para β_2 , de modo geral não há uma interpretação prática para esse parâmetro, sendo este uma constante de integração, β_3 representa a velocidade de crescimento ou taxa de maturidade do animal, β_4 é o parâmetro que determina a localização do ponto de inflexão variável e ε_i representa o efeito do erro aleatório.

Para a estimação dos parâmetros destes modelos foi usado método dos quadrados mínimos ordinários através do processo iterativo de Gauss-Newton, em que dadas as estimativas iniciais dos parâmetros e a forma da função $f(X, \theta)$, o PROC NLIN do software SAS® (SAS, 2003) ajusta o modelo $f(X, \theta) + \varepsilon$ para os dados (x_i, y_i) , $i = 1, \dots, n$ e fornece as estimativas de mínimos quadrados θ por métodos iterativos.

3.4.1 Método dos mínimos quadrados ordinários

Seja o modelo de regressão não-linear:

$$y_i = f(x_i, \theta) + \varepsilon_i, i = 1, 2, \dots, n$$

em que:

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, f(x, \boldsymbol{\theta}) = \begin{bmatrix} f(x, \theta_1) \\ f(x, \theta_2) \\ \vdots \\ f(x, \theta_n) \end{bmatrix} e \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

A função de mínimos quadrados para um modelo não-linear é:

$$SQE(\boldsymbol{\theta}) = \sum_{i=1}^n [y_i - f(x_i, \hat{\boldsymbol{\theta}})]^2$$

Na notação vetorial:

$$SQE(\boldsymbol{\theta}) = \|y - f(\hat{\boldsymbol{\theta}})\|^2$$
$$SQE(\boldsymbol{\theta}) = [y - f(\hat{\boldsymbol{\theta}})]^T [y - f(\hat{\boldsymbol{\theta}})]$$

em que:

- SQE é a soma dos quadrados dos erros aleatórios que deverá ser minimizada por $\boldsymbol{\theta}$;
- $\|y - f(\boldsymbol{\theta})\|$ é a norma euclidiana de $y - f(\boldsymbol{\theta})$

Desta forma o estimador ($\hat{\boldsymbol{\theta}}$) de mínimos quadrados de $\boldsymbol{\theta}$ é obtido através do mínimo da soma de quadrados residuais (SQE), descrita acima.

Como no caso de modelo linear, o estimador de σ^2 é:

$$\hat{\sigma}^2 = \frac{SQE}{n - p}$$

em que:

n é o número de observações;

p é o número de parâmetros.

Segundo Regazzi (2003) no caso de modelo linear os estimadores dos parâmetros tem propriedades ótimas como não tendenciosidade e variância mínima. No caso não-linear não podem ser feitas afirmações gerais sobre as propriedades dos estimadores de mínimos quadrados,

exceto para grandes amostras (chamados resultados assintóticos). Nesse caso, as propriedades como variância mínima e não tendenciosidade, por exemplo, são obtidas através do limite quando tem-se amostra de tamanho suficientemente grande. Desta forma, apenas procedimentos aproximados para os testes estatísticos e intervalos de confiança são aplicados.

Para obtenção dos estimadores de mínimos quadrados através do processo de derivação tem-se as seguintes notações de diferenciação matricial:

$f(\boldsymbol{\theta})$ é uma função vetor coluna $n \times 1$ de um argumento p dimensional $\boldsymbol{\theta}$, dada por:

$$f(\boldsymbol{\theta}) = \begin{bmatrix} f_1(\boldsymbol{\theta}) \\ f_2(\boldsymbol{\theta}) \\ \vdots \\ f_n(\boldsymbol{\theta}) \end{bmatrix}$$

$F(\boldsymbol{\theta})$ é matriz jacobiana de $f(\boldsymbol{\theta})$, dada por:

$$F(\boldsymbol{\theta}) = \frac{\partial f(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{bmatrix} \frac{\partial f_1(\boldsymbol{\theta})}{\partial \theta_1} & \frac{\partial f_1(\boldsymbol{\theta})}{\partial \theta_2} & \dots & \frac{\partial f_1(\boldsymbol{\theta})}{\partial \theta_p} \\ \frac{\partial f_2(\boldsymbol{\theta})}{\partial \theta_1} & \frac{\partial f_2(\boldsymbol{\theta})}{\partial \theta_2} & \dots & \frac{\partial f_2(\boldsymbol{\theta})}{\partial \theta_p} \\ \vdots & \vdots & \dots & \vdots \\ \frac{\partial f_n(\boldsymbol{\theta})}{\partial \theta_1} & \frac{\partial f_n(\boldsymbol{\theta})}{\partial \theta_2} & \dots & \frac{\partial f_n(\boldsymbol{\theta})}{\partial \theta_p} \end{bmatrix}_p$$

Dessa forma, o estimador de mínimos quadrados $\hat{\boldsymbol{\theta}}$ satisfaz a equação $\frac{SQE}{\partial \boldsymbol{\theta}^T} = \phi$ representa a minimização em estudo, sendo:

$$\frac{QMR}{\partial \boldsymbol{\theta}^T} = \frac{\partial}{\partial \boldsymbol{\theta}^T} [y - f(\boldsymbol{\theta})]^T [y - f(\boldsymbol{\theta})] = 2[y - f(\boldsymbol{\theta})]^T \left(\frac{-\partial f(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T} \right) = -2[y - f(\boldsymbol{\theta})]^T F(\boldsymbol{\theta})$$

Então tem-se:

$$-2[y - f(\hat{\theta})]^T F(\hat{\theta}) = \phi$$

$$-2F^T(\hat{\theta})[y - f(\hat{\theta})] = \phi$$

$$F^T(\hat{\theta})[y - f(\hat{\theta})] = \phi$$

Portanto, o sistema de equações normais (SEN), para o modelo não-linear é dado por:

$$\begin{bmatrix} \frac{\partial f_1(\hat{\theta})}{\partial \hat{\theta}_1} & \frac{\partial f_2(\hat{\theta})}{\partial \hat{\theta}_1} & \dots & \frac{\partial f_n(\hat{\theta})}{\partial \hat{\theta}_1} \\ \frac{\partial f_1(\hat{\theta})}{\partial \hat{\theta}_2} & \frac{\partial f_2(\hat{\theta})}{\partial \hat{\theta}_2} & \dots & \frac{\partial f_n(\hat{\theta})}{\partial \hat{\theta}_2} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_1(\hat{\theta})}{\partial \hat{\theta}_p} & \frac{\partial f_2(\hat{\theta})}{\partial \hat{\theta}_p} & \dots & \frac{\partial f_n(\hat{\theta})}{\partial \hat{\theta}_p} \end{bmatrix}_n \cdot \left(\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} - \begin{bmatrix} f_1(\hat{\theta}) \\ f_2(\hat{\theta}) \\ \vdots \\ f_n(\hat{\theta}) \end{bmatrix} \right) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Estas equações do SEN acima são não-lineares nos estimadores e, em geral, não existe uma solução explícita. Sendo assim, a solução para o sistema deve ser obtida por meio de processos iterativos.

3.4.2 Processos iterativos

Uma maneira de resolver problemas de mínimos quadrados não-lineares é através de procedimentos numéricos por meio de métodos iterativos no qual é determinada a solução do SEN (Sistema de equações normais).

Um método iterativo amplamente utilizado para regressão não-linear é o Método de Gauss-Newton, no qual é feita a linearização da função não-linear.

Seja $y_i = f(x_i, \theta) + \varepsilon_i$ o modelo não-linear e $\hat{\theta}_0$ um valor tal que $F^T(\hat{\theta}_0)[y - f(\hat{\theta}_0)] \approx 0$. Por expansão de série de Taylor de primeira ordem pelo ponto $\hat{\theta}_0$ tem-se: $f(\hat{\theta}) \approx f(\hat{\theta}_0) + F(\hat{\theta}_0)(\theta - \hat{\theta}_0)$.

Sendo assim, aplicando a aproximação anterior em $F^T(\hat{\theta})[y - f(\hat{\theta})] \approx \phi$ obtém-se $F^T(\hat{\theta})[y - f(\hat{\theta}_0) - F(\hat{\theta}_0)(\hat{\theta} - \hat{\theta}_0)] \approx \phi$

Ao multiplicar $[F^T(\hat{\theta})]^{-1}$ à esquerda de ambos os lados tem-se:

$$y - f(\hat{\theta}_0) - F(\hat{\theta}_0)\hat{\theta} + F(\hat{\theta}_0)\hat{\theta}_0 \approx \phi, \text{ dessa forma: } F(\hat{\theta}_0)\hat{\theta} \approx F(\hat{\theta}_0)\hat{\theta}_0 + [y - f(\hat{\theta}_0)]$$

agora multiplicando por $[F(\hat{\theta}_0)]^{-1}$ a esquerda ambos os lados tem-se:

$$\hat{\theta} \approx \hat{\theta}_0 + [F(\hat{\theta}_0)]^{-1}[y - f(\hat{\theta}_0)], \text{ agora tomando } \hat{\theta}_0 = \hat{\theta}_t \text{ e } \hat{\theta} = \hat{\theta}_{t+1} \text{ obtém-se}$$

$$\hat{\theta}_{t+1} \approx \hat{\theta}_t + [F(\hat{\theta}_t)]^{-1}[y - f(\hat{\theta}_t)] \text{ que é a expressão para a } t\text{-ésima iteração do}$$

processo iterativo de Gauss-Newton.

Isto gera um processo iterativo para determinar $\hat{\theta}$ onde se pretende que $SQE(\hat{\theta}_1)$ seja menor que $SQE(\hat{\theta}_2)$, com isso $\hat{\theta}_1$ pode ser um estimador melhor que $\hat{\theta}_2$.

O processo iterativo de Gauss-Newton continua até a convergência, isto é, até que:

$$\left| \frac{(\hat{\theta}_{j,t+1} - \hat{\theta}_{j,t})}{\hat{\theta}_{j,t}} \right| < \delta, j = 1, 2, \dots, p$$

Onde δ é algum número pequeno em que a cada iteração o $SQE(\hat{\theta}_t)$ pode ser avaliado para garantir que foi obtida uma redução neste valor, com relação ao anterior.

Após a convergência do procedimento para um vetor $\hat{\theta}$ de estimativas, podemos obter o quadrado médio residual dado por:

$$QMR = \frac{\sum_{i=1}^n [y_i - f(x_i, \hat{\theta})]^2}{n - p}, \text{ que é um estimador de } \sigma^2.$$

Um problema dos métodos iterativos é determinar os valores iniciais θ_0 , dos parâmetros, utilizados para o ajustamento de modelos de regressão não-linear. A escolha de valores de θ_0 próximos ao verdadeiro valor dos parâmetros facilita a convergência do processo iterativo.

Segundo Regazzi (2003) há diversas alternativas para auxiliar na escolha satisfatória dos valores de θ_0 , como utilizar estimativas de trabalhos anteriores semelhantes. Os parâmetros geralmente têm significados biológicos que podem ser úteis na determinação dos valores iniciais. Se o

modelo admite ser linearizado, a nova forma linear pode ser utilizada para obter o valor inicial para os parâmetros.

Nesse trabalho foi analisado o diagrama de dispersão dos dados para determinar os valores iniciais dos parâmetros.

3.5 Avaliadores da qualidade de ajuste

Para definir a relação funcional entre as variáveis ajusta-se mais de um modelo e com base nos resultados dos avaliadores da qualidade de ajuste, escolhe-se o modelo que melhor representa o processo em estudo.

Como tanto as funções splines quanto os modelos não-lineares apresentam vantagens e desvantagens sob pontos de vista estatísticos, é preciso determinar qual é o mais apropriado para descrever o crescimento corporal de uma população de ovinos e para isso serão utilizados os avaliadores de qualidade de ajuste, descritos a seguir.

Os avaliadores da qualidade de ajustes de modelos utilizados foram AIC (critério de informação de Akaike), BIC (critério de informação bayesiano), Coeficiente de determinação ajustado (R^2_{aj}).

No critério de informação de Akaike e no critério de informação bayesiano foi utilizado o logaritmo da função de máxima verossimilhança apresentado a seguir.

3.5.1 Função de verossimilhança

O objetivo do método de máxima verossimilhança é encontrar os valores dos parâmetros que maximizam a função de verossimilhança conjunta para um conjunto específico de dados.

Um modelo de regressão qualquer que apresenta resíduos com distribuição normal, como por exemplo, o modelo linear simples pode ser representado da seguinte forma:

$$y_i = \beta x_i + u_i; \text{ onde } u_i \sim N(0, \sigma^2)$$

Neste caso, a função de densidade do erro u_i é dada por:

$$f(u_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-u_i^2}{2\sigma^2}\right), \text{ para } i = 1, 2, \dots, n$$

Como erro u_i tem distribuição normal com média zero e variância σ^2 , y_i também terá distribuição normal e a função de densidade de probabilidade de cada observação será dada por:

$$f(y_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(y_i - \beta x_i)^2}{2\sigma^2}\right), \text{ para } i = 1, 2, \dots, n$$

Como as n observações são independentes a função de verossimilhança conjunta é o produto das verossimilhanças individuais. A maximização da função de verossimilhança é equivalente a maximização da função log-verossimilhança, uma vez que a função logarítmica é uma função crescente, porém, a log-verossimilhança é mais fácil de lidar.

Desta forma o logaritmo natural da função de verossimilhança conjunta será:

$$\ln \prod_{i=1}^n f(y_i) = \sum_{i=1}^n \ln f(y_i) = \frac{-n}{2} \ln(2\pi) - \frac{n}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \beta x_i)^2$$

No presente estudo, ajustou-se funções spline, porém o método não muda muito, o qual se resume ao seguinte procedimento para o caso de spline linear-linear:

A equação spline linear-linear, é dada por:

$$y = a + bx, \text{ se } x \leq x_0;$$

$$y = a + bx + c(x - x_0), \text{ se } x > x_0$$

Assim, em modelos de regressão spline tem-se que assumir para algum nó x_0 que $x_m < x_0 < x_{m+1}$, para certo índice m desconhecido. Sendo assim, para a função spline apresentada anteriormente, tem-se que o

logaritmo natural da função de verossimilhança conjunta (HERNANDEZ, 2010) será:

$$L(y_i) = \frac{-n}{2} \ln(2\pi) \frac{-n}{2} \ln(\sigma^2) \frac{-1}{2\sigma^2} \sum_{i=1}^m [y_i - (a + b x_i)]^2 \frac{-1}{2\sigma^2} \sum_{i=m+1}^n [y_i - (a + bx + c(x - x_0))]^2$$

3.5.2 Critério de informação de Akaike (AIC)

O critério de informação de Akaike é desenvolvido através da função de verossimilhança para verificar o grau da qualidade do ajuste e decidir qual o modelo é o mais adequado.

É útil quando se utiliza modelos com diferentes números de parâmetros, isto é, permite selecionar um modelo com um número reduzido de parâmetros.

Nem sempre o modelo mais parametrizado é melhor de acordo com esse critério que utiliza o princípio da parcimônia.

Valores menores de AIC indicam um ajuste melhor (AKAIKE, 1974). Sua expressão é dada por:

$$AIC = -2 \loglike + 2p$$

em que:

loglike é o valor do logaritmo natural da função de verossimilhança;

p é o número de parâmetros do modelo.

3.5.3 Critério de informação bayesiano (BIC)

Da mesma forma que o AIC, utiliza o princípio da parcimônia para escolha do melhor modelo.

Menores valores de BIC (SCHWARZ, 1978) refletem um ajuste melhor. Sua expressão é dada por:

$$BIC = -2 \loglike + p \ln(n)$$

em que:

loglike é o valor do logaritmo natural da função de verossimilhança;

p é o número de parâmetros;

n é o número de observações.

3.5.4 Coeficiente de determinação ajustado (R^2_{aj})

Segundo Souza (1998), a adequação dos modelos de regressão não-lineares pode ser medida pelo quadrado do coeficiente de correlação entre os valores observados e preditos.

O coeficiente de determinação R^2 em modelos de regressão linear que possui o β_0 (intercepto) expressa a proporção da variação explicada pelo modelo, já para o caso não-linear, essa interpretação não deve ser feita e este R^2 é utilizado mais como uma estatística descritiva (REGAZZI e SILVA, 2004)

Para comparar diferentes modelos pode ser usado o coeficiente de determinação (R^2), mas quando os modelos apresentam diferentes números de parâmetros ele pode não ser um bom indicador e nesse caso o coeficiente de determinação ajustado deve ser utilizado.

O coeficiente de determinação ajustado é corrigido pelos graus de liberdade, ou seja, mede o ajuste do modelo penalizando cada parâmetro adicional usado no modelo, por isso que é usado para comparar modelos com diferentes números de parâmetros.

A expressão do R^2_{aj} ajustado é dada por:

$$R^2_{aj} = R^2 - \left(\frac{p-1}{n-p} \right) (1 - R^2)$$

em que:

p é número de parâmetros do modelo;

n é o número de observações utilizadas;

R^2 é o coeficiente de determinação dado por:

$$R^2 = \frac{\left[\sum_{i=1}^n y_i \hat{y}_i - \frac{\left(\sum_{i=1}^n y_i \right) \left(\sum_{i=1}^n \hat{y}_i \right)}{n} \right]^2}{\left[\sum_{i=1}^n y_i^2 - \frac{\left(\sum_{i=1}^n y_i \right)^2}{n} \right] \left[\sum_{i=1}^n \hat{y}_i^2 - \frac{\left(\sum_{i=1}^n \hat{y}_i \right)^2}{n} \right]}$$

y_i é o peso observado;

\hat{y}_i é o peso estimado.

3.6. Identidade de Modelos

Foi testada a igualdade dos parâmetros das funções splines ajustadas para os três grupos genéticos através do método de identidade de modelos de regressão não-linear, apresentado por Regazzi (2003) e Regazzi e Silva (2004). Foi utilizado o método da razão verossimilhança com aproximação dada pela estatística qui-quadrado (χ^2). As estimativas dos parâmetros dos modelos completos e reduzidos foram obtidas utilizando o procedimento NLIN do SAS® (SAS, 2003).

Para tal método é usado o modelo que terá os parâmetros testados acrescido de uma variável indicadora (*dummy*) que representa cada um dos três grupos genéticos.

Considere o seguinte modelo denominado completo, para a função spline linear-linear:

Para $x < x_0$:

$$\text{Se } x \geq \sum_{j=1}^3 D_j [x_{0j}] + e$$

Se $x < \sum_{j=1}^3 D_j [x_{0j}] + e$, então o modelo será dado por:

$$y_{ij} = \sum_{j=1}^3 D_j [a_j + b_j x] + e_i,$$

Para $x \geq x_0$:

Se $x \geq \sum_{j=1}^3 D_j [x_{0j}] + e$, então o modelo será dado por:

$$y_{ij} = \sum_{j=1}^3 D_j [a_j + b_j x + c_j (x - x_{0j})] + e_i,$$

Considere também o seguinte modelo reduzido, para o qual é assumido um único parâmetro a para os três grupos genéticos.

Para $x < x_0$:

Se $x < \sum_{j=1}^3 D_j [x_{0j}] + e$, então o modelo será dado por:

$$y_{ij} = \sum_{j=1}^3 D_j [a + b_j x] + e_i,$$

Para $x \geq x_0$:

Se $x \geq \sum_{j=1}^3 D_j [x_{0j}] + e$, então o modelo será dado por:

$$y_{ij} = \sum_{j=1}^3 D_j [a + b_j x + c_j (x - x_0)] + e_j,$$

em que:

- a_j, b_j, c_j são os parâmetros do modelo para cada um dos 3 grupos genéticos, sendo $j=1, 2$ e 3 ;
- x_0 é o nó, ponto de junção das duas equações;
- e_i é o termo de erro aleatório, $e_i \sim N(0, \sigma^2)$;
- D_j é uma variável indicadora, em que:

$$D_j = \begin{cases} 1, & \text{se a observação } y_{ij} \text{ pertence ao grupo } j; \\ 0, & \text{caso contrário.} \end{cases}$$

Tal modelo reduzido é utilizado no método da razão de verossimilhança (Regazzi, 2003; Regazzi e Silva, 2004), através do qual são testadas as seguintes hipóteses separadamente:

$H_0^{(1)}$: $a_1 = a_2 = a_3 = a$ versus $H_a^{(1)}$: nem todos a_i são iguais;

$H_0^{(2)}$: $b_1 = b_2 = b_3 = b$ versus $H_a^{(2)}$: nem todos b_i são iguais;

$H_0^{(3)}$: $c_1 = c_2 = c_3 = c$ versus $H_a^{(3)}$: nem todos c_i são iguais;

$H_0^{(4)}$: $x_{01} = x_{02} = x_{03} = x_0$ versus $H_a^{(4)}$: nem todos x_{0i} são iguais.

As somas de quadrados dos resíduos $\hat{\sigma}_\Omega^2$ e $\hat{\sigma}_\omega^2$ dos modelos completos e reduzidos respectivamente foram usadas para obter as estatísticas qui-quadrado para o teste $H_0^{(1)}, H_0^{(2)}, H_0^{(3)}$, em que sua expressão é dada por:

$$\chi_{calculado}^2 = -n \ln \left(\frac{\hat{\sigma}_\Omega^2}{\hat{\sigma}_{\omega_i}^2} \right)$$

em que:

$$\hat{\sigma}_\Omega^2 = \frac{SQR_\Omega}{n};$$

$$\hat{\sigma}_\omega^2 = \frac{SQR_\omega}{n}$$

A regra de decisão consiste em rejeitar H_0 a um nível de significância α se $\chi_{calculado}^2 \geq \chi_\alpha^2(v)$, onde $v = p_\Omega - p_\omega$ é o número de graus de liberdade da estatística do teste, sendo p_Ω e p_ω o número de parâmetros estimados nos modelos completo e reduzido, respectivamente.

Se H_0 for rejeitada existe pelo menos uma diferença entre as estimativas dos parâmetros dos três grupos genéticos analisados. Sendo

assim, foram feitas comparações duas a duas para verificar quais grupos genéticos são iguais ou diferentes em relação ao parâmetro a .

Para tal análise considere o modelo completo como exemplo para comparação entre os grupos genéticos 1 e 2:

Se $x < d_1(x_{01}) + d_2(x_{02})$ então o modelo é: $d_1(a_1 + b_1x) + d_2(a_2 + b_2x)$

Se $x \geq d_1(x_{01}) + d_2(x_{02})$ então o modelo é: $d_1(a_1 + b_1x + c_1(x - x_{01})) + d_2(a_2 + b_2x + c_2(x - x_{02}))$

E o modelo reduzido é dado por:

Se $x < d_1(x_{01}) + d_2(x_{02})$ então o modelo é: $d_1(a + b_1x) + d_2(a + b_2x)$

Se $x \geq d_1(x_{01}) + d_2(x_{02})$ então o modelo é: $d_1(a + b_1x + c_1(x - x_{01})) + d_2(a + b_2x + c_2(x - x_{02}))$.

4. RESULTADOS E DISCUSSÃO

4.1 Estimativas dos parâmetros das funções splines

As estimativas dos parâmetros, caracterizados por a, b, c, d e nós (x_{01} e x_{02}), (ganho de peso diário no segundo e terceiro segmento, GP_2 e GP_3 respectivamente e do peso adulto), das funções splines ajustadas aos dados de peso-idade dos grupos genéticos DMN, DSI e DRL estão apresentadas respectivamente nas Tabelas 3, 4 e 5.

Tabela 3 - Estimativas dos parâmetros das funções splines para o grupo DMN, com os respectivos erros padrão assintótico

Parâmetros	Splines		
	Linear-Linear	Linear-platô	Linear-linear-linear
a	2,9570 \pm 0,72	4,6052 \pm 0,53	3,1124 \pm 0,64
b	0,1242 \pm 0,01	0,0979 \pm 0,01	0,1206 \pm 0,01
c	-0,0529 \pm 0,01	-	-0,0820 \pm 0,03
GP_2 (kg/dia)	0,0713 \pm 0,02	-	0,0386 \pm 0,03
d	-	-	0,0654 \pm 0,04
GP_3 (kg/dia)	-	-	0,1040 \pm 0,08
x_{01} (dias)	118 \pm 17,52	246 \pm 13,00	144 \pm 14,90
Peso adulto (kg)	-	28,6983 \pm 7,48	-
x_{02} (dias)	-	-	211 \pm 26,55

$GP_2=b+c$ e $GP_3 b+c+d$ são os ganhos de peso no primeiro e segundo segmentos, respectivamente.

Tabela 4 - Estimativas dos parâmetros e das funções splines para o grupo DSI, com os respectivos erros padrão assintótico

Parâmetros	Splines		
	Linear-Linear	Linear-platô	Linear-linear-linear
a	4,4535 ± 0,68	4,4866 ± 0,69	4,5515 ± 0,67
b	0,1552 ± 0,01	0,1546 ± 0,01	0,1534 ± 0,01
c	-0,170 ± 0,01	-	-0,1353 ± 0,02
GP ₂ (kg/dia)	-0,0148 ± 0,02	-	0,0181 ± 0,02
d	-	-	-0,1724 ± 0,05
GP ₃ (kg/dia)	-	-	-0,1543 ± 0,08
x ₀₁ (dias)	162 ± 6,77	157,1 ± 5,77	156 ± 9,33
Peso adulto (kg)		28,7742 ± 0,73	
x ₀₂ (dias)	-	-	267,5 ± 12,44

GP₂=b+c e GP₃ b+c+d são os ganhos de peso no primeiro e segundo segmentos, respectivamente.

Tabela 5 - Estimativas dos parâmetros e das funções splines para o grupo DRL, com os respectivos erros padrão assintótico

Parâmetros	Splines		
	Linear-Linear	Linear-platô	Linear-linear-linear
a	4,3699 ± 0,53	4,5737 ± 0,52	3,9228 ± 1,04
b	0,1161 ± 0,01	0,1130 ± 0,01	0,1322 ± 0,03
c	-0,1112 ± 0,01	-	-0,0224 ± 0,03
GP ₂ (kg/dia)	0,0049 ± 0,02	-	0,1098 ± 0,07
d	-	-	-0,1180 ± 0,01
GP ₃ (kg/dia)	-	-	-0,0082 ± 0,09
x ₀₁ (dias)	188 ± 10,68	197,9 ± 7,48	49,0008 ± 51,24
Peso adulto (kg)	-	26,936 ± 0,56	-
x ₀₂ (dias)	-	-	202,9 ± 11,19

GP₂=b+c e GP₃ b+c+d são os ganhos de peso no primeiro e segundo segmentos, respectivamente.

Para as funções spline linear-linear, considerando os três grupos genéticos, percebe-se que os pesos ao nascer (parâmetro a) variaram de 2,95 kg a 4,45 kg (Tabelas 3, 4 e 5). Enquanto o ganho de peso diário para o primeiro segmento variou de 0,11 a 0,15 kg/dia (parâmetro b). Entretanto, o ganho de peso diário, para o segundo segmento (GP_2), quando os animais estão com idade entre 118 e 188 dias (parâmetro x_{01}), caiu para 0,0713 kg/dia no grupo DMN e 0,0049 kg/dia para o DRL (Tabelas 3,4 e 5). Nessas idades os animais apresentam um desenvolvimento menor que nas idades iniciais, ou seja, passam de uma fase de crescimento acelerado para uma inibitória (DUARTE, 1975).

No entanto, para o grupo DSI (Tabela 4) o GP_2 foi negativo (-0,0148 kg/dia). Esse fato pode ser devido à presença de *outlier* ou perda de peso do animal.

Percebe-se que, de forma geral, as estimativas dos parâmetros da função spline linear-platô foram de acordo com a realidade. Os valores estimados para o peso ao nascimento (parâmetro a) variaram de 4,48 kg a 4,6 kg. O ganho de peso diário estimado ficou entre 0,09 e 0,15 kg/dia (Tabelas 3, 4 e 5), sendo o grupo DSI o que apresentou maior estimativa (parâmetro b).

Os pesos à maturidade dos animais (peso no platô) foram de 26,93 a 28,79 kg (Tabelas 3, 4 e 5). Estas estimativas concordam com aquelas obtidas por Silveira (2010) (28,83 a 46,83 kg para o peso adulto dos animais destes mesmos grupos genéticos). O grupo genético DMN atingiu o peso adulto com 246 dias, enquanto o grupo DSI atingiu com 157 dias e o DRL com 198 dias (Tabelas 3, 4, e 5). Assim, verifica-se que o grupo DSI apresenta maior precocidade. Este resultado era esperado uma vez que a raça Santa Inês é mais precoce que as raças Morada Nova e Rabo Largo.

Os valores estimados para os pesos adultos, obtido através do platô, representaram bem a realidade, uma vez que de acordo com Albuquerque e Meyer (2001), em geral, o crescimento animal atinge um platô à maturidade.

Para a função spline linear-linear-linear tem-se que o peso ao nascimento dos animais variou de 3,11 a 4,55 kg. Enquanto o ganho de peso diário ficou entre 0,12 e 0,15kg/dia para o primeiro segmento (GP_1) e para o segundo esses valores ficaram entre 0,0181 a 0,1098 kg/dia (GP_2).

Exceto para o grupo genético DMN, no terceiro segmento, o ganho de peso diário (GP_3) apresentou valores negativos -0,1543 e -0,0082 para os grupos DSI e DRL, respectivamente (Tabelas 4 e 5).

As Figuras 1, 2, 3, 4, 5, 6, 7, 8 e 9 ilustram os gráficos de dispersão das funções splines linear-linear, linear-platô, linear-linear-linear estimadas para os grupos genéticos DMN, DSI e DRL.

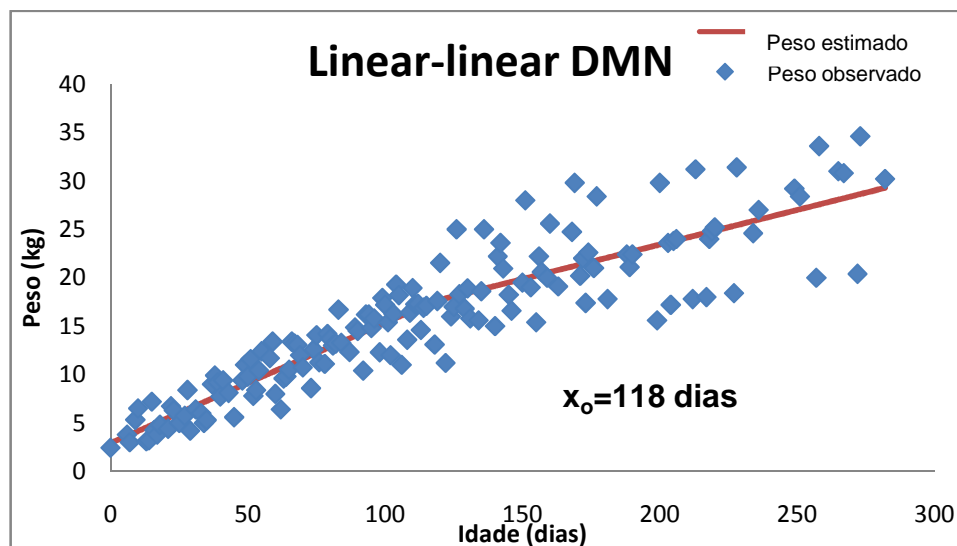


Figura 1 - Função spline linear-linear para o grupo genético DMN

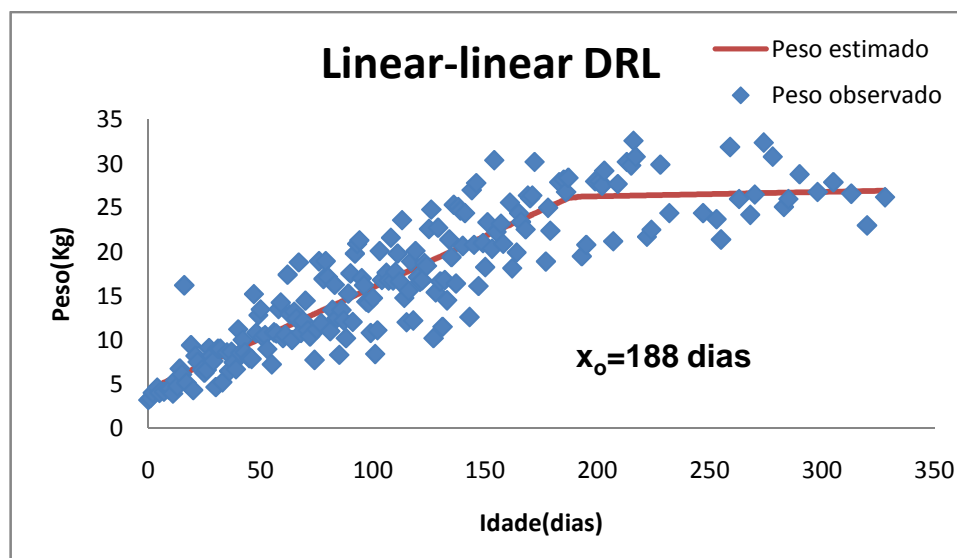


Figura 2 - Função spline linear-linear para o grupo genético DRL

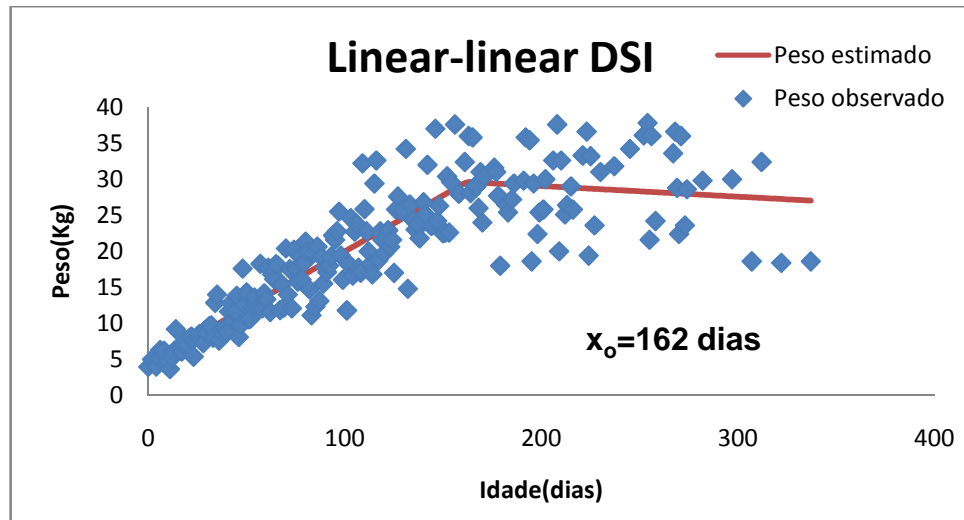


Figura 3 - Função spline linear-linear para o grupo genético DSI

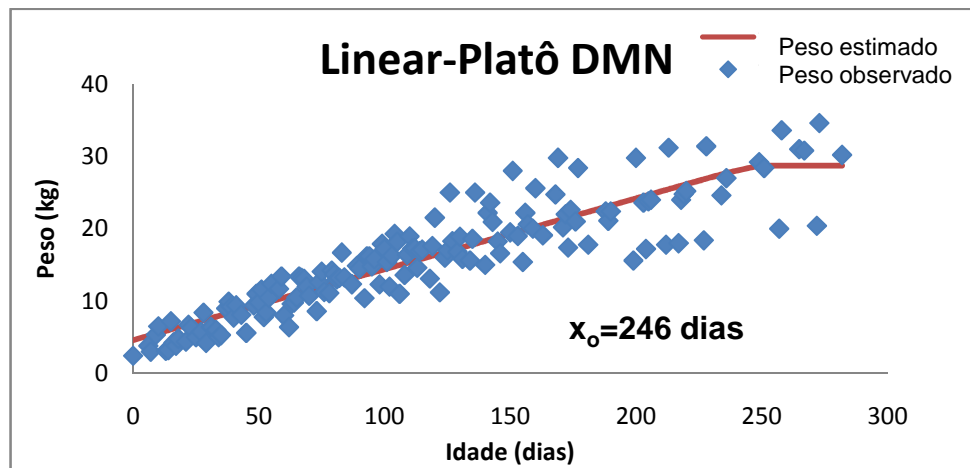


Figura 4 - Função spline linear-platô para o grupo genético DMN

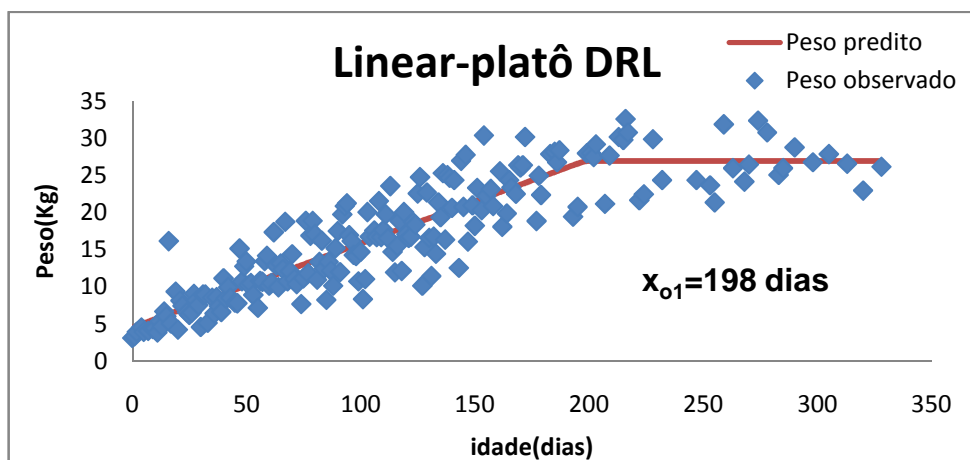


Figura 5 - Função spline linear- platô para o grupo genético DR

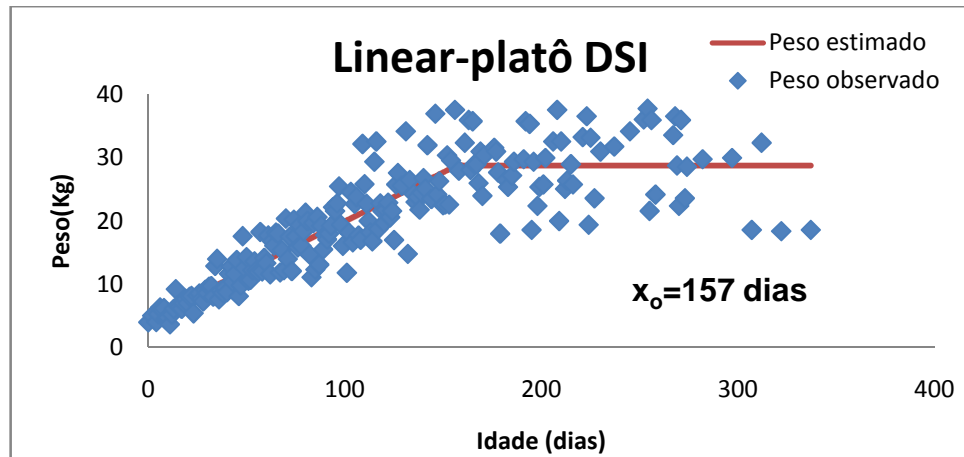


Figura 6 - Função spline linear-platô para o grupo genético DSI

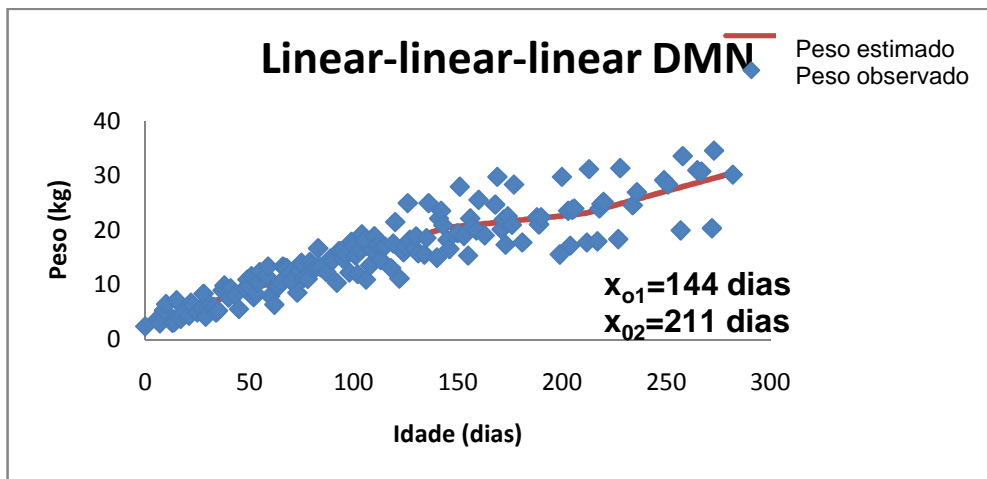


Figura 7 - Função spline linear-linear-linear para o grupo DMN

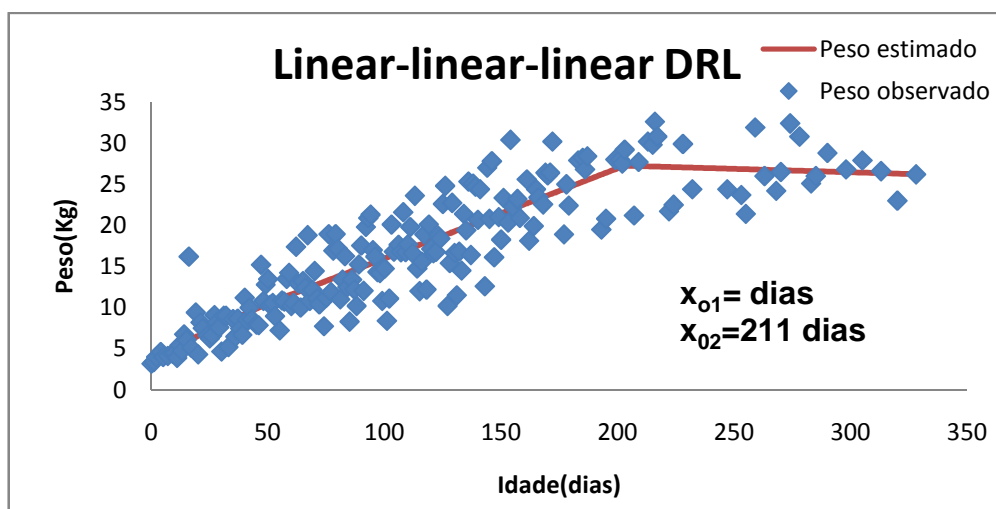


Figura 8 - Função spline linear-linear-linear para o grupo DRL

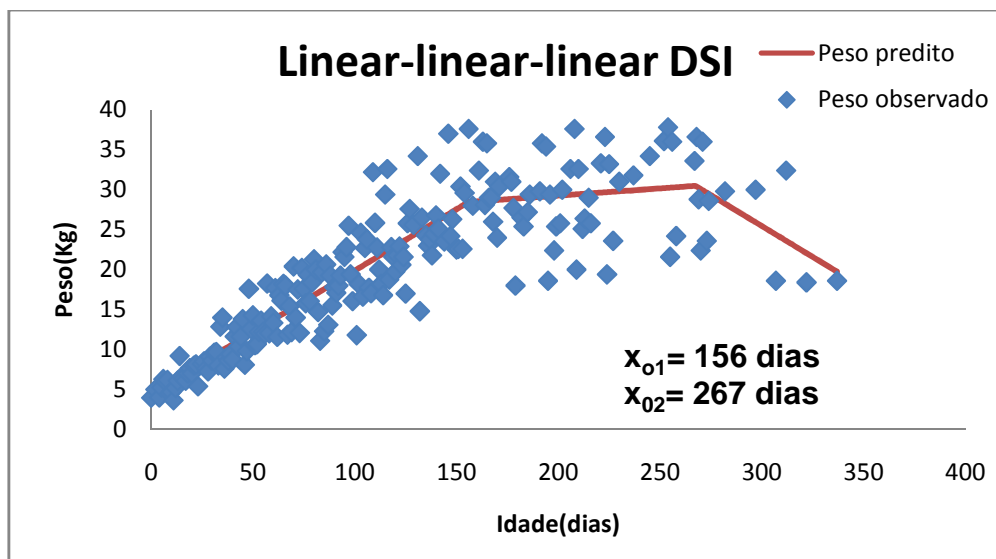


Figura 9 - Função spline linear-linear-linear para o grupo DSI

4.2 Estimativas dos parâmetros para modelos não-lineares usuais

As estimativas dos parâmetros dos modelos não-lineares Richards, Brody e von Bertalanffy, ajustados aos dados de peso-idade dos grupos genéticos DMN, DRL e DSI estão apresentados nas Tabelas 6, 7 e 8 respectivamente.

Tabela 6 - Estimativas dos parâmetros da regressão não-linear de Richards para os três grupos genéticos

Parâmetros	DMN	DRL	DSI
β_1	$28,9893 \pm 3,8701$	$27,5782 \pm 1,0203$	$29,2863 \pm 0,8002$
β_2	$-3,9965 \pm 100,1$	$3,3100 \pm 1,7729$	$2,9402 \pm 1,7371$
β_3	$0,0113 \pm 0,0060$	$0,0254 \pm 0,0084$	$0,0310 \pm 0,00962$
β_4	$0,00816 \pm 0,8110$	$2,1017 \pm 1,2146$	$1,7777 \pm 1,1183$

Tabela 7 - Estimativas dos parâmetros da regressão não-linear de Brody para os três grupos genéticos

Parâmetros	DMN	DRL	DSI
β_1	41,0072 ± 6,2117	34,4741 ± 2,6543	32,4900 ± 1,4017
β_2	0,9423 ± 0,0161	0,9187 ± 0,0204	0,9546 ± 0,0329
β_3	0,00404 ± 0,0010	0,00576 ± 0,0009	0,0097 ± 0,00120

Tabela 8 - Estimativas dos parâmetros da regressão não-linear de von Bertalanffy para os três grupos genéticos

Parâmetros	DMN	DRL	DSI
β_1	32,1820 ± 2,4268	30,6808 ± 1,4791	30,8015 ± 0,9688
β_2	0,5261 ± 0,0225	0,5015 ± 0,0224	0,5366 ± 0,0340
β_3	0,00862 ± 0,0012	0,00994 ± 0,0099	0,0149 ± 0,00148

Em geral, percebe-se que as estimativas dos parâmetros estão de acordo com a realidade. Os valores obtidos para a taxa de maturidade (β_3) e peso adulto (β_1) (Tabela 6, 7 e 8) concordam com aqueles obtidos por Silveira (2010) que também utilizou dados de crescimento destes mesmos grupos genéticos. Nota-se também que os pesos adultos estimados pela função spline linear-platô (Tabelas 3, 4 e 5) foram condizentes aos estimados pelos modelos não-lineares usuais. Através das funções spline estimou 28,69 (Tabela 3) para o peso adulto do grupo DMN, 28,77 para o grupo DSI (Tabela 4) e 26,93 para o grupo DRL (Tabela 4).

Silveira (2010) obteve, respectivamente para os grupos DMN, DSI e DRL os valores $\beta_1 = 17,85; 27,46; 25,60, \beta_3 = 0,08; 0,98; 1,06$ para o modelo Richards, $\beta_1 = 27,61; 36,25; 32,34, \beta_3 = 0,01; 0,011; 0,008$ para o modelo Brody e $\beta_1 = 34,47; 38,70; 34,20, \beta_3 = 0,012; 0,013; 0,011$ para o modelo von Bertalanffy.

As Figuras 10, 11 e 12 ilustram os gráficos dos ajustes dos modelos não-lineares Richards, Brody e von Bertalanffy para o grupo genético DMN.

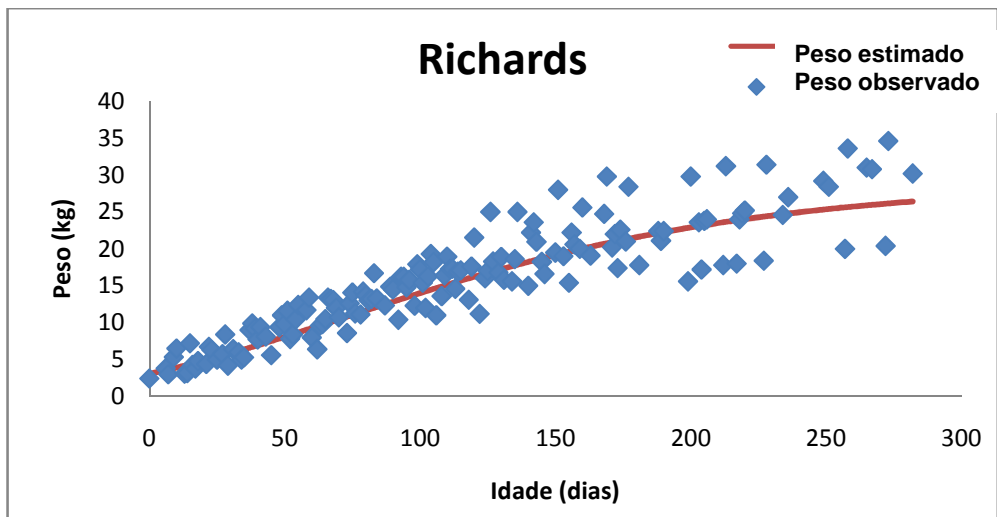


Figura 10 - Modelo não-linear Richards para o grupo genético DMN

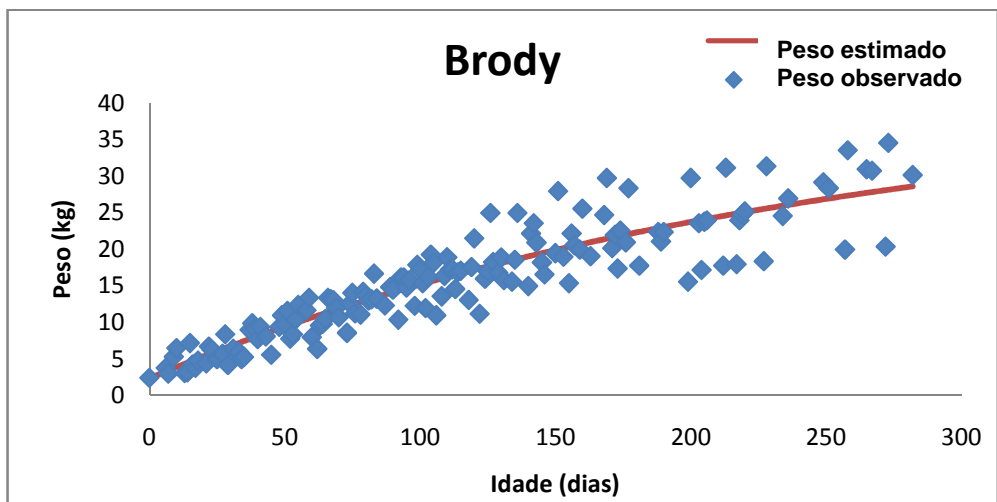


Figura 11 - Modelo não-linear Brody para o grupo DMN

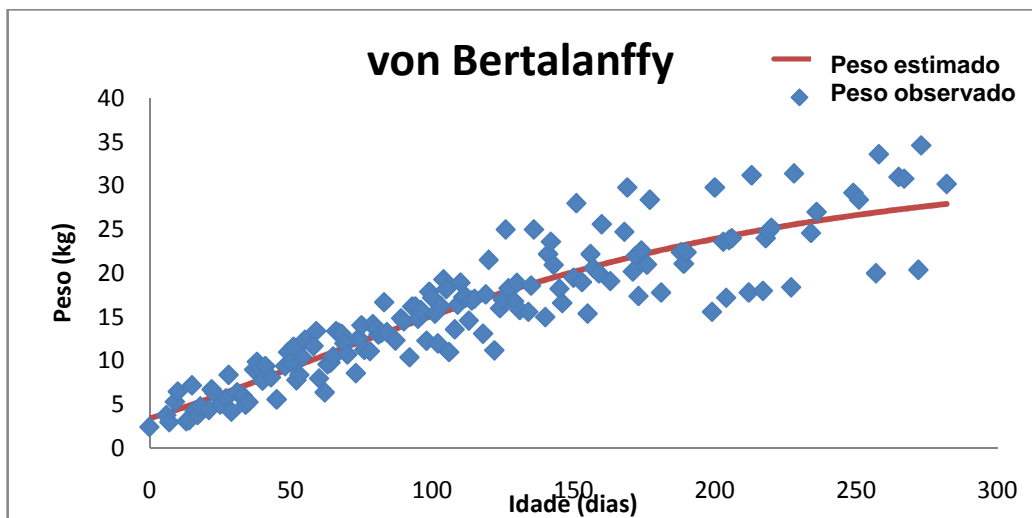


Figura 12 - Modelo não-linear von Bertalanffy para o grupo DMN

4.3 Avaliadores da qualidade de ajuste para regressão splines

Nas Tabelas 9, 10 e 11 são apresentados, respectivamente os avaliadores da qualidade de ajuste AIC, BIC e R_{aj}^2 para os grupos genéticos DMN, DRL e DSI.

Tabela 9 - Resultados do avaliador da qualidade de ajuste, critério de informação de Akaike (AIC), para as funções splines

Splines	Nº de parâmetros	Grupos genéticos		
		DMN	DSI	DRL
Linear-linear	4	773,6879	1185,2198	1028,5943
Linear-platô	3	649,8101	1184,9916	1026,5980
Linear-linear-linear	6	774,4627	1177,4217	1031,5581

Tabela 10 - Resultados do avaliador da qualidade de ajuste, critério de informação bayesiano (BIC), para as regressões splines

Splines	Nº de parâmetros	Grupos genéticos		
		DMN	DSI	DRL
Linear-linear	4	785,7570	1198,5700	1041,6863
Linear-platô	3	658,8620	1195,0042	1036,4170
Linear-linear-linear	6	792,5664	1197,4469	1051,1961

Tabela 11 - Resultados do avaliador da qualidade de ajuste, coeficiente de determinação ajustado (R_{aj}^2 %), para as regressões splines

Splines	Nº de parâmetros	Grupos genéticos		
		DMN	DSI	DRL
Linear-linear	4	82,80	79,13	81,23
Linear-platô	3	81,24	79,05	81,33
Linear-linear-linear	6	91,08	90,09	89,50

De acordo com os resultados apresentados nas Tabelas 9, 10 e 11, para o grupo DMN a spline linear-platô apresentou melhor ajuste, uma vez que as estimativas do AIC e BIC apresentaram os menores valores para esse modelo, além de R_{aj}^2 superior a 80%.

Através das estimativas do R_{aj}^2 e AIC, utilizadas na indicação do melhor modelo, a spline linear-linear-linear se ajustou melhor aos dados de peso-idade para o grupo genético DSI.

Entretanto, a spline linear-platô apresentou menor valor para o BIC e os resultados dos demais avaliadores da qualidade de ajuste foram muito próximos da spline linear-linear-linear, com isso a spline linear-platô pode ser considerada a que apresentou melhor ajuste, também para o grupo DSI.

Ao comparar as estimativas dos avaliadores da qualidade do ajuste pode-se verificar que para o grupo genético DRL a spline linear-platô foi a que melhor se ajustou aos dados, uma vez para esse modelo o AIC e BIC apresentaram menores valores e R_{aj}^2 de 81,34%.

4.4 Avaliadores da qualidade de ajuste para modelos não-lineares

Nas Tabelas 12, 13 e 14 são apresentados, os avaliadores da qualidade de ajuste AIC, BIC e R_{aj}^2 para os grupos genéticos DMN, DRL e DSI, respectivamente.

Tabela 12 - Estimativas dos avaliadores da qualidade de ajuste dos modelos não-lineares usuais para o grupo DMN

Modelos	Nº de parâmetros	AIC	BIC	R_{aj}^2 %
Brody	3	772,5767	781,6285	82,82
von Bertalanffy	3	773,7905	782,8424	82,68
Richards	4	794,4532	806,5223	82,36

Tabela 13 - Estimativas dos avaliadores da qualidade de ajuste dos modelos não-lineares usuais para o grupo DRL

Modelos	Nº de parâmetros	AIC	BIC	R_{aj}^2
Brody	3	1043,3010	1053,1200	79,6693
von Bertalanffy	3	1037,7567	1047,5757	80,2410
Richards	4	1034,5972	1047,6892	80,6558

Tabela 14 - Estimativas dos avaliadores da qualidade de ajuste dos modelos não-lineares usuais para o grupo DSI

Modelos	Nº de parâmetros	AIC	BIC	R_{aj}^2
Brody	3	1206,9625	1216,9751	76,7248
von Bertalanffy	3	1197,7906	1207,8032	77,7343
Richards	4	1192,8267	1206,1769	78,3561

Para o grupo genético DMN o modelo não-linear Brody foi o que melhor descreve a relação entre as variáveis peso-idade, uma vez que para todos os avaliadores da qualidade de ajuste esse modelo é o que apresentou melhor desempenho (Tabela 12).

De acordo com as estimativas dos avaliadores da qualidade do ajuste, apresentadas na Tabela 13, para o grupo genético DRL, o modelo não-linear Richards foi que melhor se ajustou aos dados. Para esse modelo todos avaliadores da qualidade de ajuste apresentaram um resultado igual ou superior com relação aos demais modelos.

Com relação ao grupo genético DSI, o modelo Richards também foi o que apresentou melhor ajuste, apresentando valores menores para AIC e BIC e maior para o R_{aj}^2 (Tabela 14).

4.5 Funções splines e modelos não-lineares usuais

Na Tabela 15, são apresentados os avaliadores da qualidade de ajuste AIC, BIC e R_{aj}^2 para as funções splines e os modelos não-lineares usuais que apresentaram melhor ajuste para os grupos genéticos DMN, DRL e DSI.

Tabela 15 - Estimativas dos avaliadores da qualidade de ajuste para funções splines e modelos não-lineares usuais que apresentaram melhor ajuste para os grupos genéticos DMN, DRL e DSI

Grupos	Avaliadores	Modelos	Splines
		Brody	Linear-platô
DMN	AIC	772,57	649,81
	BIC	781,62	658,86
	$R_{aj}^2(\%)$	82,82	81,24
	Nº de parâmetros	3	3
DRL		Richards	Linear-platô
	AIC	1034,59	1026,59
	BIC	1047,68	1036,41
	$R_{aj}^2(\%)$	80,65	81,33
	Nº de parâmetros	4	3
DSI		Richards	Linear- platô
	AIC	1192,82	1184,99
	BIC	1206,17	1195,00
	$R_{aj}^2(\%)$	78,35	79,05
	Nº de parâmetros	4	3

De acordo com as estimativas obtidas na Tabela 15, verificou-se que para todos os grupos genéticos as funções splines Linear-platô ajustaram-se

melhor do que os modelos não-lineares usuais. Apresentando valores altos para o R^2_{aj} e menores para AIC e BIC.

A grande vantagem das funções splines é poder ajustar o crescimento do animal de forma diferente em diferentes intervalos de idade. Além disso, são de fáceis manipulações, apresentam interpretação direta para alguns parâmetros.

4.6 Identidade de modelos

Nas Tabelas 16 e 17, são apresentados os resultados dos testes de razão de verossimilhança com aproximação de qui-quadrado com as respectivas hipóteses e valor de p, para avaliação da identidade de funções splines ajustadas para os três grupos genéticos.

Tabela 16 - Testes de identidade de parâmetros do modelo, considerando spline Linear-linear, para os grupos genéticos (DRL=1, DMN=2 e DSI =3)

Hipóteses	n	SQR $_{\Omega}$	SQR $_{\omega}$	$\chi^2_{Calculado}$	Valor de p
		Completo	Reduzido		
H $_0$: x $_{01}$ =x $_{02}$ =x $_{03}$	554	7045,2	7139,5	7,3661*	0,0251
H $_0$: x $_{01}$ =x $_{02}$	346	3549,2	3618,3	6,6715*	0,0097
H $_0$: x $_{02}$ =x $_{03}$	359	4904,4	4920	1,1401 ^{ns}	0,2856
H $_0$: x $_{01}$ =x $_{03}$	403	5636,8	5710,4	5,2279*	0,0222
H $_0$: a $_1$ =a $_2$ =a $_3$	554	7045,2	7068,1	1,7978 ^{ns}	0,4070
H $_0$: b $_1$ =b $_2$ =b $_3$	554	7045,2	7342,7	22,913*	0,00001
H $_0$: b $_1$ =b $_2$	346	3549,2	3546,5	0 ^{ns}	1
H $_0$: b $_2$ =b $_3$	359	4904,4	4933	2,0874 ^{ns}	0,1485
H $_0$: b $_1$ =b $_3$	403	5636,8	6060,2	29,187*	6,57x10 ⁻⁸
H $_0^{(4)}$: c $_1$ =c $_2$ =c $_3$	554	7045,2	7195,6	11,702*	0,0028
H $_0$: c $_1$ =c $_2$	346	3549,2	3639,1	8,6540*	0,0032
H $_0$: c $_2$ =c $_3$	359	4904,4	5020,0	8,3636*	0,0038
H $_0$: c $_1$ =c $_3$	403	5636,8	5679,1	3,0129 ^{ns}	0,0826

SQR $_{\Omega}$ é a soma de quadrado do modelo completo; SQR $_{\omega}$ é a soma de quadrado do modelo reduzido; * Significativo ao nível de 5% de probabilidade; ^{ns} não significativo ao nível de 5% de probabilidade.

De acordo com os resultados da estatística do teste da razão de verossimilhança (Tabela 16), o parâmetro x_0 no grupo DMN (118 dias) diferiu do grupo DRL (188 dias) que também diferiu do grupo DSI (162 dias). Entretanto, os parâmetros para os grupos genéticos DMN e DSI não diferiram estatisticamente, dessa forma estes grupos (DMN e DSI) passam para uma fase de ganho de peso diário menor com idade inferior ao grupo DRL.

Com relação ao peso ao nascimento (parâmetro a) os três grupos genéticos não diferiram estatisticamente.

O grupo DRL diferiu do grupo DSI quanto ao parâmetro ganho de peso diário (parâmetro b), para o primeiro segmento, e este último pode ser considerado estatisticamente igual ao grupo DMN, sendo esse igual ao DSI (Tabela 16).

Tendo em vista o sinal negativo para a estimativa do parâmetro c conclui-se que houve decréscimo no ganho de peso diário no segundo segmento ($b+c$). De acordo com a Tabela 16 percebe-se que o decréscimo no ganho de peso do grupo DRL (-0,11) diferiu do grupo DMN (-0,05) que também diferiu do grupo DSI (-0,17).

As estimativas obtidas para os grupos genéticos DRL e DSI mostraram-se estatisticamente iguais, pode-se inferir que para o grupo DMN a taxa de crescimento apresentou um decréscimo inferior aos demais grupos para o segundo segmento.

Tabela 17 - Testes de identidade de parâmetros de modelos considerando a spline Linear-platô, para os grupos genéticos (DRL=1, DMN=2, DSI)

Hipóteses	n	SQR _Ω	SQR _ω	χ ² _{Calculado}	Valor-p
		Completo	Reduzido		
H ₀ : X ₀₁ =X ₀₂ =X ₀₃	554	7214,1	9789,1	169,09*	1,90x10 ⁻⁷
H ₀ : X ₀₁ =X ₀₂	346	3733	3734,9	0,1760 ^{ns}	0,6747
H ₀ : X ₀₂ =X ₀₃	359	5116,5	5233,3	8,1031*	0,0044
H ₀ : X ₀₁ =X ₀₃	403	5668,4	5888,4	15,345*	8,95x10 ⁻⁵
H ₀ : a ₁ =a ₂ =a ₃	554	7214,1	7068,1	11,326*	0,0034
H ₀ : a ₁ =a ₂	346	3733	3740,6	0,70 ^{ns}	0,4015
H ₀ : a ₂ =a ₃	359	5116,5	5126,8	0,7219 ^{ns}	0,3954
H ₀ : a ₁ =a ₃	403	5668,4	5668,4	0 ^{ns}	1
H ₀ : b ₁ =b ₂ =b ₃	554	7214,1	7624,1	30,622*	2,23x10 ⁻⁷
H ₀ : b ₁ =b ₂	346	3733	3737,9	0,4538 ^{ns}	0,5005
H ₀ : b ₂ =b ₃	359	5116,5	5338,5	15,24*	9,42x10 ⁻⁵
H ₀ : b ₁ =b ₃	403	5668,4	5944,4	19,159*	1,20x10 ⁻⁵

SQR_Ω é a soma de quadrado do modelo completo; SQR_ω é a soma de quadrado do modelo reduzido; * Significativo ao nível de 5% de probabilidade; ^{ns} não significativo ao nível de 5% de probabilidade.

Tendo em vista os resultados da estatística do teste da razão de verossimilhança, mostrados na Tabela 17, para spline linear-platô, verificou-se que o ponto x₀ em que o grupo DMN (246 dias, Tabela 3) atingiu o platô diferiu do grupo DSI (157 dias, Tabela 4) e este último também diferiu estatisticamente do grupo DRL (198 dias, Tabela 5). No entanto, os grupos DMN e DRL não diferiram no ponto de mudança, quanto à idade ao atingir o platô. Portanto, o grupo DSI atingiu o platô, ou seja, o peso adulto com idade inferior aos demais grupos, sendo mais precoce.

Quanto ao ganho de peso diário (parâmetro *b*), o grupo genético DRL (0,11 kg/dia) difere do DSI (0,15 kg/dia), e este último diferiu do grupo DMN (0,09 kg/dia), no entanto os grupos DMN e DRL não diferiram estatisticamente. Dessa forma o grupo DSI foi o que apresentou maior ganho de peso diário e maior precocidade, sendo o grupo mais recomendado para produção de carne.

5. CONCLUSÕES

As funções splines mostraram-se mais adequadas do que os modelos não-lineares usuais para descrever curvas de crescimento dos três grupos genéticos de ovinos cruzados. Verificou-se que todos os avaliadores da qualidade de ajuste utilizados indicaram melhores ajustes com utilização de funções splines.

A função spline linear-platô apresentou melhor ajuste em relação às splines linear-linear e linear-linear-linear. Além de possuir facilidade na interpretação dos parâmetros, principalmente, em relação à idade em que ocorre mudança na fase de crescimento ($nó$) e no ganho de peso diário (parâmetro b).

De acordo com testes de identidade de modelos, o grupo genético Dorper x Santa Inês foi o que apresentou maior precocidade e maior ganho de peso diário, sendo o mais indicado para exploração de carne.

6. REFERÊNCIAS BIBLIOGRÁFICAS

ALBUQUERQUE, L.G.; MEYER, K. Estimates of covariance functions for growth from birth to 630 days of age in Nelore cattle. *Journal of Animal Science*, v.79, p.2776-2789, 2001.

AKAIKE, H. A new look at the statistical model identification. *IEEE Transaction on Automatic Control*, v.19, p.716-723, 1974.

BRODY, S. *Bioenergetics and Growth*. Rheinhold Publishing, New York. 1945.

DUARTE, F. A. M. Estudo da curva de crescimento de animais da raça Nelore (*Bos taurus indicus*) através de cinco modelos estocásticos. 1975. 284 f. Tese (Livre-Docência) - Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 1975.

ELIAS, M. A. Análise de curvas de crescimento de vacas das raças Nelore, Guzerá e Gir. 1998. 128 f. Dissertação (Mestrado em Zootecnia) - Escola Superior de Agricultura de "Luiz de Queiroz", Piracicaba, 1998.

FRANÇOIS, P. Desempenho, características de carcaça e a utilização da carne de ovelhas de descarte terminadas em pastagem cultivada na elaboração de embutido fermentado. 2009. 85f. Dissertação (Curso de Mestrado do Programa de Pós-Graduação em Zootecnia) Universidade Federal de Santa Maria, Santa Maria, RS, 2009.

GONÇALVES, H.C. Fatores genéticos e de meio em algumas características produtivas e reprodutivas de caprinos. Viçosa: Tese (doutorado em Zootecnia) – Universidade Federal de Viçosa, 1996. 141 p.

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.

HERNANDEZ, E.L. Parameter Estimation in Linear-Linear Segmented Regression. 2010. 93f. Dissertação (Mestrado) faculty of Brigham Young University.

HUISMAN, A.E.; VEERKAMP, R.F.; ARENDONK, J.A.M. Genetics parameters for various random regression models to describe the weight data of pigs. *Journal of Animal Science*, v.80, p.575-582, 2002.

IBGE. Produção da pecuária municipal rebanho bovino brasileiro: efetivo de rebanhos 2008. Disponível em: <<http://www.ibge.gov.br>>. Acesso em: 20 jan. 2011.

KEELE, L.J. *Semiparametric regression Estatística for the science*. John Wiley and Sons, 1^o edição, 2008.

LOBO, R.N.B.; VILLELA, L.C.V.; LOBO, A.M.B.O.; PASSOS, J.R.de.S.; OLIVEIRA, A.A.de. Parâmetros genéticos de características estimadas da curva de crescimento de ovinos da raça Santa Inês. *Revista Brasileira de Zootecnia*, v.35, n.3,p.1012-1019, 2006 (supl.).

MALHADO, C.H.M.; CARNEIRO, P.L.S.; AFFONSO, P.R.A.M.; SOUZA Jr., A.A.O.; SARMENTO, J.L.R. Growth curves in Dorper sheep crossed with the local Brazilian breeds, Morada Nova, Rabo Largo, and Santa Inês. *Small Ruminant Research*, v.84, p.16-21, Jun 2009.

OLIVEIRA, T. S. Ovinocultura de corte no Brasil: Abordagem mercadológica, *Revista da Terra*, junho de 2010.

PAULSON, D.S. *Handbook of regression and modeling – Application for the clinical and pharmaceutical industries*. Chapman & Hall / CRC Biostatistics Series, 2007.

REGAZZI, A.J. Teste para verificar a igualdade de parâmetros e a identidade de modelos de regressão não-linear. *Revista Ceres*, v.50, n.287, p.9-26, 2003.

REGAZZI, A.J.; SILVA, C.H.O. Teste para verificar a igualdade de parâmetros e a identidade de modelos de regressão não-linear. I. Dados no delineamento inteiramente casualizado. *Revista Matemática e, São Paulo*, v.22, n.3, p.33-45, 2004.

RICHARDS, F.J. A flexible growth function for empirical use. *Journal of Experimental Botany, Oxford*, v.10, p.290-300, 1959.

ROBBINS, K.R.; MISZTAL, I.; BERTRAND, J.K. A practical longitudinal model for evaluating growth in Gelbvieh cattle. *J. Anim. Sci.*, v.83, p.29-33, 2005.

SAS Institute Inc. *Statistical Analysis System user's guide*. Version 9.1 ed. Cary: SAS Institute, USA, 2003, licenciado pela Universidade Federal de Viçosa.

SCHENKEL, F.S. Utilização de polinômios segmentados na pesquisa zootécnica (considerações teóricas e práticas). Universidade Federal do Rio Grande do Sul – Faculdade de Agronomia – Departamento de Zootecnia, Seminário em Zootecnia, Porto Alegre – Rio grande do Sul, 1989.

SCHWARZ, G. Estimating the dimension of a model. *Annals of Statistics*, v.6, p.461-464, 1978.

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. *Ciência e Agrotecnologia, Lavras*, v.25, n.5, p.1195-1205, set./out. 2001.

SILVEIRA, F.G. Classificação multivariada de modelos de crescimento para grupos genéticos de ovinos de corte, 2010, 59f. Dissertação (Mestrado) Universidade Federal de Viçosa– Departamento de Estatística, Viçosa – Minas Gerais.

SORIO, A. Carne ovina e caprina: produção e consumo no Brasil e nas Américas, último acesso em 25/07/2010 disponível em <http://www.neapl.sedetec.se.gov.br/modules/news/article.php>

SOUSA, W. H.; LEITE, P. R. M. Ovinos de corte: a raça Dorper. João Pessoa: EMEPA-PB, 76 p, 2000.

SOUZA, J. C.; BIANCHINI SOBRINHO, E. Estimativa do peso de bovinos de corte, aos 24 meses, da raça Nelore, usando curvas de crescimento. Revista da Sociedade Brasileira de Zootecnia, Viçosa, v. 23, n. 1, p. 85-91, jan./fev. 1994.

SOUZA, G.S. Introdução aos modelos de regressão linear e não-linear. Brasília: Embrapa-SPI, 1998. 505p.

VARGA, G. Teste de modelos estatísticos para a estrutura a termo no Brasil. Revista Brasileira de Economia. 2009, vol.63, n.4, pp. 361-394. ISSN 0034-7140.

VIANA, J. G. A. Panorama Geral da Ovinocultura no Mundo e no Brasil. Revista Ovinos, Nº12, Porto Alegre, Ano 4, Março de 2008.

VON BERTALANFFY, L. Quantitative laws for metabolism and growth. Q. Rev. Biol. 32:217–231, 1957.

APÊNDICES

APÊNDICE A – Rotinas utilizadas no software SAS® considerando o ajuste da função spline linear-platô aos dados de peso-idade dos três grupos genéticos DMN.

```
data teste;
  input x y raca $;

  i_drl=(raca='drl');i_dmn=(raca='dmn');i_dsi=(raca='dsi');

datalines;
0 3.185 drl
1 3.4 drl
2 4 drl
... ..
337 18.6 dsi
;
proc print;run;

/*modelo completo*/

PROC NLIN DATA =teste maxiter=1000 CONVERGEPARM=0.001
/*1=drl, 2=dmn, 3=dsi*/
PARMS a1= 5.5 b1= 0.09 x01= 270
      a2= 5.5 b2= 0.09 x02= 270
      a3= 3 b3= 0.6 x03= 110;
if x LE (i_drl*x01 + i_dmn*x02 + i_dsi*x03) then
MODEL y = i_drl*(a1 + b1*x) + i_dmn*(a2 + b2*x) +
i_dsi*(a3 + b3*x);
ELSE
MODEL y = i_drl*(a1 +b1*x01) + i_dmn*(a2+b2*x02) +
i_dsi*(a3 +b3*x03);
run;

/*modelo incompleto p/ x0*/

PROC NLIN DATA =teste maxiter=1000 CONVERGEPARM=0.001;
/*1=drl, 2=dmn, 3=dsi*/
PARMS a1= 3.5 b1= 0.16 x0= 100
      a2= 3 b2= 0.6
      a3= 3 b3= 0.6 ;
if x LE x0 then
MODEL y = i_drl*(a1 + b1*x) + i_dmn*(a2 + b2*x) +
i_dsi*(a3 + b3*x);
ELSE
```



```

MODEL y = i_drl*(a1 + b1*x0) + i_dmn*(a2 + b2*x0) +
i_dsi*(a3 + b3*x);
;
run;

/*modelo incompleto p/ a*/

PROC NLIN DATA =teste maxiter=1000 CONVERGEPARM=0.001;
PARMS a= 3.5 b1= 0.16 c1= -0.5 x01= 100
      b2= 0.6 c2= -0.5 x02= 110
      b3= 0.6 c3= -0.5 x03= 110;
if x LE (i_drl*x01 + i_dmn*x02 + i_dsi*x03) then
MODEL y = i_drl*(a + b1*x) + i_dmn*(a + b2*x) + i_dsi*(a
+ b3*x);
ELSE
MODEL y = i_drl*(a - c1* x01 + (b1+c1)*x) + i_dmn*(a -
c2* x02 + (b2+c2)*x) + i_dsi*(a - c3* x03 + (b3+c3)*x);
;
run;

/*modelo incompleto p/ b*/

PROC NLIN DATA =teste maxiter=1000 CONVERGEPARM=0.001
PARMS a1= 3.5 b= 0.16 x01= 100
      a2= 3 x02= 110
      a3= 3 x03= 110;
if x LE (i_drl*x01 + i_dmn*x02 + i_dsi*x03) then
MODEL y = i_drl*(a1 + b*x) + i_dmn*(a2 + b*x) +
i_dsi*(a3 + b*x);
ELSE
MODEL y = i_drl*(a1 +b*x01) + i_dmn*(a2+b*x02) +
i_dsi*(a3 +b*x03);
run;

/* comparando 2x2 drl x dmn 1=drl, 2=dmn, 3=dsi*/

data drlxdmn; set teste; if raca ne "dsi";
proc print;run;

/* modelo completo */

PROC NLIN DATA =drlxdmn maxiter=1000 CONVERGEPARM=0.001 ;
PARMS a1= 3.5 b1= 0.16 x01= 100
      a2= 3 b2= 0.6 x02= 180;

if x LE (i_drl*x01 + i_dmn*x02 ) then
MODEL y = i_drl*(a1 + b1*x) + i_dmn*(a2 + b2*x) ;
ELSE
MODEL y = i_drl*(a1 + b1*x01) + i_dmn*(a2 + b2*x02);
run;

```

```

/* modelo p a */

PROC NLIN DATA =drlxdmn maxiter=1000 CONVERGEPARM=0.001 ;
PARMS a= 3.5 b1= 0.16 x01= 100
      b2= 0.6 x02= 180;

if x LE (i_drl*x01 + i_dmn*x02 ) then

MODEL y = i_drl*(a + b1*x) + i_dmn*(a + b2*x) ;
ELSE
MODEL y = i_drl*(a + b1*x01) + i_dmn*(a + b2*x02);
run;

/* modelo p/ x0 */

PROC NLIN DATA =drlxdmn maxiter=1000 CONVERGEPARM=0.001 ;
PARMS a1= 3.5 b1= 0.16 c1= -0.5 x0= 150
      a2= 3 b2= 0.6 c2= -0.5;

if x LE x0 then
MODEL y = i_drl*(a1 + b1*x) + i_dmn*(a2 + b2*x) ;
ELSE
MODEL y = i_drl*(a1 + b1*x0) + i_dmn*(a2 + b2*x0);
run;

/* modelo p/ b */

PROC NLIN DATA =drlxdmn maxiter=1000 CONVERGEPARM=0.001 ;
PARMS a1= 3.5 b= 0.3 x01= 100
      a2= 3 x02= 180;

if x LE (i_drl*x01 + i_dmn*x02 ) then

MODEL y = i_drl*(a1 + b*x) + i_dmn*(a2 + b*x) ;
ELSE
MODEL y = i_drl*(a1 + b*x01) + i_dmn*(a2 + b*x02);
run;

/* comparando dmn x dsi 1=drl, 2=dmn, 3=dsi*/

data dmnXdsi; set teste; if raca ne "drl"; /* criando o
conjunto de dados dmn x dsi*/
proc print;run;

/* modelo completo*/

PROC NLIN DATA =dmnXdsi maxiter=1000 CONVERGEPARM=0.001 ;
PARMS a2= 3.5 b2= 0.16 x02= 100
      a3= 3 b3= 0.6 x03= 180;

```

```

if x LE ( i_dmn*x02 +i_dsi*x03) then
MODEL y = i_dmn*(a2 + b2*x)+i_dsi*(a3 + b3*x) ;

ELSE
MODEL y =i_dmn*(a2 + b2*x02)+ i_dsi*(a3 + b3*x03) ;
run;

/* modelo p/ x0 */

PROC NLIN DATA =dmnXdsi maxiter=1000 CONVERGEPARM=0.001 ;
PARMS a2= 3.5 b2= 0.16 x0= 100
      a3= 3 b3= 0.6 ;

if x LE x0 then
MODEL y = i_dmn*(a2 + b2*x)+i_dsi*(a3 + b3*x) ;
ELSE
MODEL y =i_dmn*(a2 + b2*x0)+ i_dsi*(a3 + b3*x0) ;
run;

/* modelo p a*/

PROC NLIN DATA =dmnXdsi maxiter=1000 CONVERGEPARM=0.001 ;
PARMS a= 3.5 b2= 0.16 x02= 100
      b3= 0.6 x03= 180;

if x LE ( i_dmn*x02 +i_dsi*x03) then
MODEL y = i_dmn*(a + b2*x)+i_dsi*(a + b3*x) ;

ELSE
MODEL y =i_dmn*(a + b2*x02)+ i_dsi*(a + b3*x03) ;
run;

/* modelo p b*/

PROC NLIN DATA =dmnXdsi maxiter=1000 CONVERGEPARM=0.001 ;
PARMS a2= 3.5 b= 0.16 x02= 100
      a3= 3 x03= 180;

if x LE ( i_dmn*x02 +i_dsi*x03) then
MODEL y = i_dmn*(a2 + b*x)+i_dsi*(a3 + b*x) ;

ELSE
MODEL y =i_dmn*(a2 + b*x02)+ i_dsi*(a3 + b*x03) ;
run;

/* comparando drl x dsi*/

data drlXdsi; set teste; if raca ne "dmn"; /* criando o
conjunto de dados drl x dsi 1=drl, 2=dmn, 3=dsi*/
proc print;run;

```

```

/* modelo completo*/

PROC NLIN DATA =drlXdsi maxiter=1000 CONVERGEPARM=0.001;
PARMS a1= 3.5 b1= 0.16 x01= 100
      a3= 3 b3= 0.6 x03= 110;

if x LE (i_drl*x01 + i_dsi*x03 ) then
MODEL y = i_drl*(a1 + b1*x) + i_dsi*(a3 + b3*x) ;
ELSE
MODEL y = i_drl*(a1 + b1*x01) + i_dsi*(a3 + b3*x03);
run;

/* modelo p/ x0 */

PROC NLIN DATA =drlXdsi maxiter=1000 CONVERGEPARM=0.001;
PARMS a1= 3.5 b1= 0.16 x0= 100
      a3= 3 b3= 0.6 ;

if x LE x0 then

MODEL y = i_drl*(a1 + b1*x) + i_dsi*(a3 + b3*x) ;
ELSE
MODEL y = i_drl*(a1 + b1*x0) + i_dsi*(a3 + b3*x0);
run;

/* modelo p/ incompleto a*/

PROC NLIN DATA =drlXdsi maxiter=1000 CONVERGEPARM=0.001;
PARMS a= 3.5 b1= 0.16 x01= 100
      b3= 0.6 x03= 110;

if x LE (i_drl*x01 + i_dsi*x03 ) then
MODEL y = i_drl*(a + b1*x) + i_dsi*(a + b3*x) ;
ELSE
MODEL y = i_drl*(a + b1*x01) + i_dsi*(a + b3*x03);
run;

/* modelo incompleto p/ b*/

PROC NLIN DATA =drlXdsi maxiter=1000 CONVERGEPARM=0.001;
PARMS a1= 3.5 b= 0.16 x01= 100
      a3= 3 x03= 110;

if x LE (i_drl*x01 + i_dsi*x03 ) then
MODEL y = i_drl*(a1 + b*x) + i_dsi*(a3 + b*x) ;
ELSE
MODEL y = i_drl*(a1 + b*x01) + i_dsi*(a3 + b*x03);
run;

```