THALITA KELEN LEAL DO PRADO

REGRESSÃO NÃO LINEAR MULTIVARIADA NO CRESCIMENTO DO COCO VARIEDADE ANÃ VERDE

LAVRAS - MG

2016

THALITA KELEN LEAL DO PRADO

REGRESSÃO NÃO LINEAR MULTIVARIADA NO CRESCIMENTO DO COCO VARIEDADE ANÃ VERDE

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, área de concentração em Estatística e Experimentação Agropecuária, para a obtenção do título de Doutor.

Orientador

Dr. Joel Augusto Muniz

Coorientadora

Dra. Taciana Villela Savian

LAVRAS - MG

2016

Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).

Prado, Thalita Kelen Leal do.
Regressão não linear multivariada no crescimento do coco variedade anã verde / Thalita Kelen Leal do Prado. – Lavras : UFLA, 2016.
62 p. : il.

Tese(doutorado)–Universidade Federal de Lavras, 2016.
Orientador: Joel Augusto Muniz.
Bibliografia.

1. Modelo de crescimento. 2. Análise de regressão multivariada. 3. Mínimos quadrados. I. Universidade Federal de Lavras. II. Título.

THALITA KELEN LEAL DO PRADO

REGRESSÃO NÃO LINEAR MULTIVARIADA NO CRESCIMENTO DO COCO VARIEDADE ANÃ VERDE

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, área de concentração em Estatística e Experimentação Agropecuária, para a obtenção do título de Doutor.

APROVADA em 11 de fevereiro de 2016.

Dra. Taciana Villela Savian	ESALQ
Dr. Daniel Furtado Ferreira	UFLA
Dr. Augusto Ramalho de Morais	UFLA
Dr. Ângelo Albérico Alvarenga	EPAMIG

Dr. Joel Augusto Muniz
Orientador

LAVRAS - MG

2016

Aos meus pais, Jair e Maria, pelo amor incondicional.

Aos meus irmãos, Junior e Suziléa,

A minha cunhada Ariane.

Ao meu esposo Alexandre, pelo amor e compreensão.

Dedico.

AGRADECIMENTOS

A Deus, em primeiro lugar, por me guiar, dar sabedoria e permitir mais esta conquista em minha vida. *Soli Deo gloria*!

Ao professor Dr. Joel Augusto Muniz, pela oportunidade, incentivo e orientação no decorrer deste curso.

À professora Dra. Taciana Villela Savian, pela disponibilidade, incentivo, materiais compartilhados e colaborações na pesquisa.

Aos professores do Departamento de Ciências Exatas (DEX), pelos ensinamentos transmitidos que colaboraram para o meu crescimento.

À Universidade Federal de Lavras (UFLA) e ao Departamento de Ciências Exatas (DEX), pela oportunidade de realizar o doutorado.

Ao Fundo de Amparo à Pesquisa do estado de Minas Gerais - FAPEMIG e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES pela concessão da bolsa de estudos.

A Josi, ex-secretária do Departamento de Ciências Exatas (DEX) e a Nádia, secretária atual, pela atenção, boa vontade e simpatia sempre prestadas.

Aos colegas e amigos Adriele, Tales, Sílvio, Thaís Destéfani, Elayne, Wederson Leandro, Alexandre Adão, Thallis, Eugênia, Suzana, Andréia, Vania, Lidia e família, Maria Eduarda, Alex que de perto ou longe contribuíram em muitos momentos de aprendizado e descontração, sempre serei grata a vocês.

Em especial, à minha grande amiga Adriana, pelo companheirismo, amor e carinho e ao seu esposo, grande amigo Pedro Sorato *in memoriam*, por tantos momentos de convivência agradável. Sempre serei grata pelo enorme apoio e amizade sincera.

A toda a minha família, principalmente aos meus pais Jair e Maria, pelo amor, paciência, dedicação, atenção, orações e incentivo para que eu não desistisse. Sem o apoio e o encorajamento de vocês eu não teria chegado até aqui, serei eternamente grata por tudo que fizeram por mim.

Ao meu esposo Alexandre, pelo amor, companheirismo e apoio em momentos difíceis durante a trajetória do doutorado.

RESUMO

O coqueiro anão verde é um dos principais recursos vegetais da humanidade cuja exploração tem evoluído na maioria dos estados brasileiros, em decorrência do interesse comercial na água do coco para consumo *in natura* e uso na indústria de envasamento, ocupando espaço no vultoso mercado de refrigerantes. Nos plantios comerciais destinados ao mercado de água no Brasil, predomina a variedade anã, em virtude da sua boa performance, em termos de rendimento e qualidade da água de coco. Os modelos de regressão não linear logístico e gompertz na descrição do crescimento de plantas e frutos têm sido utilizados. Conduziu-se o presente estudo, com o objetivo de avaliar o ajuste dos modelos não lineares multivariados logístico (LL), gompertz (GG) e híbridos (GL e LG), por meio do método de mínimos quadrados, considerando estrutura de erros independentes e autorregressiva de segunda ordem para os resíduos, a fim de obter as estimativas dos parâmetros. Os modelos multivariados foram ajustados a dados de crescimento de fruto de coqueiro anão verde, diâmetro externo longitudinal e transversal. A escolha do melhor modelo foi feita por meio do critério de informação de Akaike corrigido, coeficiente de determinação ajustado e quadrado médio residual, sendo que todos os modelos apresentaram um bom ajuste e o modelo multivariado gompertz GG, considerando estrutura autorregressiva de segunda ordem, foi o mais adequado para ajustar os dados experimentais, resultando em estimativas coerentes com os apresentados na literatura. Os procedimentos de ajuste dos modelos de regressão foram realizados pelo programa computacional Statistical Analysis System SAS.

Palavras-chave: Modelo de crescimento. Análise de regressão multivariada. Mínimos Quadrados.

ABSTRACT

The green dwarf coconut palm is a major plant resource of humanity whose operation has evolved in most brazilian states because the commercial interest in coconut juice for fresh consumption and use in potting industry, taking up space on bulky soft drinks market. In commercial plantations for the water market in Brazil dominates the dwarf variety, because of its good performance in terms of yield and quality of coconut water. The nonlinear regression models logistic and gompertz growth in the description of plants and fruits have been used. The objective of this study was to evaluate the fit of nonlinear logistic multivariate models (LL), gompertz (GG) and hybrid (GL and LG) by the method of least squares, structure basis of independent and autoregressive errors of the second order for the residuals, to obtain estimates of the parameters. The multivariate models were adjusted growth data of green dwarf coconut fruit, longitudinal and transverse outer diameter. Choosing the best model was made using the Akaike information criterion corrected coefficient of determination adjusted and residual mean square and all the models showed a good fit and the multivariate model gompertz GG, considering auto regressive structure of the second order, were more adequate to fit the experimental data, resulting in estimates consistent with those reported in the literature. The adjustment procedures regression models were performed by the computer program Statistical Analysis System SAS.

Keywords: Growth model. Multivariate regression analysis. Least squares.

SUMÁRIO

1 INTRODUÇÃO

O coqueiro *Cocos nucifera L.* da variedade anã é composto das cultivares amarela, verde, vermelho de Camarões e vermelho da Malásia, sendo a cultivar verde, a principal demanda para plantio no Brasil. O coqueiro anão constitui-se na variedade de coqueiro mais utilizada comercialmente no Brasil, para produção de água de coco, com qualidade sensorial superior às demais cultivares; sendo empregada também na agroindústria de alimentos e/ou do fruto seco *in natura*, com produtividade estimada de polpa nos plantios tecnificados, acima de 8 ton/ha. Tendo em vista a informação anteriormente citada, essa variedade pode se constituir em alternativa promissora para os produtores de coco seco, pois, além de se tornar uma variedade de maior utilidade comercial, reduzirá o "déficit" de produção de polpa atualmente observado nos plantios com as cultivares de coqueiro híbrido e gigante (ARAGÃO, 2007).

A análise do crescimento e desenvolvimento de frutos é fundamental para descrever as mudanças ocorridas nas características físicas e componentes químicos, possibilitando conhecer a dinâmica envolvida, auxiliando nas estratégias adotadas na gestão do processo de produção e colheita. A avaliação dos processos de desenvolvimento de um fruto permite estabelecer as bases para definir o ponto mais adequado de colheita e estratégias para sua conservação, visando ao aumento da vida útil pós-colheita.

A dinâmica do crescimento pode ser auxiliada por meio de fórmulas ou modelos, sendo o primeiro passo utilizado para quantificar a produção vegetal, pois independente das dificuldades inerentes ao conhecimento sobre a complexidade que envolve o crescimento de plantas e frutos, a análise quantitativa do crescimento é uma ferramenta acessível e bastante precisa para avaliar o desenvolvimento vegetal. Em geral, existem muitos modelos estatísticos que quantificam a produção vegetal bem como descrevem curvas de crescimento de plantas e suas partes (caule, folha, fruto, sementes e raízes).

Tendo em vista a importância do estudo das curvas de crescimento, vários tipos de modelos estatísticos são usados para estudar os processos de crescimento envolvidos no sistema de produção vegetal, sendo os modelos de regressão lineares e não lineares comumente utilizados. Os modelos lineares são bastante utilizados visto serem facilmente implementados nas análises estatísticas, porém há situações que não têm sido adequados para descrever fenômenos biológicos. Já, os modelos não lineares, estes fornecem informações e estimativas úteis, particularmente se forem baseados em parâmetros que permitem interpretação biológica, como obtenção da taxa de crescimento, peso e altura máximos.

Os trabalhos que utilizam modelos de regressão para o estudo de fenômenos de crescimento, em geral, empregam análises para o caso univariado, ou seja, modelos em que possuem variável resposta única. No entanto, ao analisar os fenômenos aleatórios, pode-se identificar que todos, sejam eles culturais ou naturais, envolvem um grande número de variáveis. Dessa forma, torna-se interessante utilizar modelos de regressão multivariados, já que é comum medir mais do que uma variável resposta na unidade experimental e ao considerar, simultaneamente, todas as variáveis respostas, as estimativas dos parâmetros do modelo serão mais precisas.

Dentre os modelos não lineares que podem ser utilizados para expressar o crescimento de frutos, por meio da análise multivariada, estão os modelos gompertz e logístico. E, visto que, na literatura consultada são escassas as informações sobre o ajuste de modelos não lineares, por meio da análise multivariada, pretendeu-se, com este trabalho, estudar o ajuste dos modelos gompertz, logísticos e híbridos, para o caso multivariado, considerando a estrutura independente e autorregressiva nos resíduos, a dados de crescimento de frutos de coqueiro da variedade anã verde.

2 REFERENCIAL TEÓRICO

2.1 A cultura do coco anão

A fruticultura é uma das atividades mais importantes da agricultura brasileira, respondendo por 25 % do valor da produção agrícola nacional. Além disso, houve um aumento na área e ampliou suas fronteiras em direção à região nordeste, onde condições de luminosidade, umidade relativa e temperatura são mais favoráveis do que nas regiões Sul e Sudeste (LACERDA; LACERDA; ASSIS, 2004).

O gênero Cocos é constituído apenas pela espécie *Cocos nucifera L.*, a qual é composta de algumas variedades, entre as quais as mais importantes são: *Typica* (Var. Gigante) e *Nana* (Var. Anã). Os híbridos de coqueiro mais utilizados são resultantes dos cruzamentos entre essas variedades. O nome *Cocos nucifera*, segundo Benassi (2006), infere a planta que produz nozes com aspecto de cabeça, pois a palavra "coco" se origina do português que significa cabeça e "nucifera", do Latim "nucifer-a-um", que produz nozes.

Dentre os principais países produtores de coco do mundo, a variedade anã não tem utilidade comercial, sendo empregada apenas nos programas de cruzamento intervarietal e para fins ornamentais. Já, no Brasil, além do programa de melhoramento, o coco anão é muito utilizado para extração da água de coco, com produção média em torno de 300 mL/fruto e com qualidade sensorial superior as outras cultivares de coqueiro (LOIOLA, 2009).

Segundo o Sindicato dos Produtores de Coco (SINDCOCO), em torno de 70%, 20% e 10% dos plantios de coqueiro no país, são formados pelas cultivares gigante, anã e híbrida, respectivamente (ARAGÃO, 2007). O coqueiro da variedade anã originou-se, provavelmente, de uma mutação ocorrida na variedade Gigante e apresenta frutos com pouca polpa e água muito saborosa destinada ao consumo, seja sob a forma “in natura” ou processadas em agroindústrias (MACIEL et al., 2009).

O coqueiro da variedade anã apresenta crescimento vegetativo lento, é precoce, iniciando a produção em média com dois a três anos após o plantio. A planta chega a atingir 10 a 12m de altura e tem vida útil em torno de 30 a 40 anos. Apresenta estipe delgado, folhas numerosas, porém curtas, produz um grande número de pequenos frutos (150 a 200 frutos/planta/ano), é mais sensível ao ataque de pragas, como ácaro, e doenças foliares. Em geral, apresenta maiores exigências de clima e solo do que a variedade Gigante (ARAGÃO, 2007).

O coqueiro anão tem crescimento e produção contínua durante todo o ano, desde que as condições do clima sejam ideais. Em locais onde a evapotranspiração é elevada e a distribuição das chuvas é irregular, ocorrem déficits hídricos sazonais que afetam o desenvolvimento e a produção (HOLANDA et al., 2007). Da mesma forma como ocorre em outras culturas, no cultivo de coco irrigado a utilização da técnica da fertigação também tende a aumentar e de acordo com Ferreira Neto et al. (2011) o número de folhas, diâmetro de copa, altura e circunferência do estipe, os quais foram favorecidos pela aplicação de nitrogênio e potássio.

De acordo com Bitencourt e Pedroti (2008), o coqueiro anão verde constitui-se na variedade de coqueiro mais utilizada comercialmente no Brasil, para a produção de água de coco, com qualidade sensorial superior às das demais cultivares. Os frutos destinados ao consumo “in natura” de sua água devem ser colhidos, principalmente, entre o sexto e o sétimo mês, após a abertura natural da inflorescência, independente da cultivar considerada. Nessa idade, ocorrem os maiores valores para peso de fruto, produção de água, teores de frutose, glicose, grau brix e sais minerais, principalmente potássio, os quais conferem melhor sabor à água de coco. Os frutos colhidos precocemente apresentam água menos doce, e naqueles colhidos tardiamente ocorrem aumentos nos teores de gordura, ocasionando um sabor rançoso à água de coco.

Aragão (2007) recomenda efetuar uma mistura da água dos frutos colhidos nas idades de 5 a 8 meses para o coco anão destinado ao uso agroindustrial. Já, os frutos secos para produção agroindustrial de alimentos ou para uso culinário, devem ser colhidos entre 11 a 12 meses de idade, isto é, frutos maduros.

O ponto ideal de colheita do fruto está associado a uma série de indicadores morfológicos relacionados à idade, ao tamanho do fruto e à presença de determinadas substâncias na água. Depende também de determinadas propriedades químicas e sensoriais, ligadas aos aspectos nutritivos, alimentares e de saúde humana. De acordo com Farias et al. (2011), o ponto de colheita e o manejo após a colheita dos frutos dependem do padrão fisiológico de cada fruto. No processo de desenvolvimento, o fruto de coco inicialmente cresce mais em comprimento do que em largura, ou seja, com o diâmetro transversal menor que o diâmetro longitudinal, atingindo o tamanho e peso máximos aos seis meses de idade (ARAGÃO, 2007).

O conhecimento das mudanças das características físicas durante o desenvolvimento dos frutos de coqueiro da variedade anã verde, dentre outras variedades, serve como ferramenta para as indústrias que tèm como atividade econômica a exploração desses frutos, pois essas informações servem para identificação de padrões de qualidade e identidade, bem como também apontar para a exploração comercial de novas cultivares, gerando, então, uma nova alternativa de produção com a utilização de cultivares ainda não difundidas comercialmente (MACIEL et al., 2009).

2.2 Análise de crescimento

O termo análise de crescimento, no contexto vegetal, refere-se a um conjunto útil de métodos quantitativos que descrevem e interpretam a performance de todo o sistema vegetal, seja ele cultivado em condições naturais, seminaturais ou controladas (HUNT, 1978). No estudo de crescimento de plantas, Urchei, Rodrigues e Stone (2000) afirmam que, para avaliar os efeitos de sistemas de manejo sobre as plantas, a análise de crescimento é fundamental, pois descreve as mudanças na produção vegetal em função do tempo, o que não é possível com o simples registro do rendimento.

O estudo de crescimento na área vegetal possui uma vasta aplicação na pesquisa agropecuária, pois possibilita ao pesquisador comparar o comportamento de uma planta ou de suas partes de interesse, diante das diversas situações experimentais utilizadas.

Cardoso et al. (2006) destacam que uma forma de maximizar a eficiência de plantas em captação dos recursos limitantes (nutrientes, luz, água entre outros) ao crescimento, desenvolvimento e produção é manejar adequadamente a cultura, com a utilização de espaçamento e/ou arranjo espacial eficientes, uma população de plantas adequada, cultivar adaptada ao local específico, entre outros fatores. A análise de crescimento, neste caso, entra como uma ferramenta importante que avalia a eficiência desses manejos, sendo bastante utilizada por pesquisadores das Ciências Agrárias. Peixoto et al. (2006) afirmam que na mensuração das alterações no crescimento vegetal, o acúmulo de matéria seca é a medida mais significante, uma vez que resulta da associação de vários outros componentes.

Segundo Cavalini et al. (2006), o ponto de colheita é um dos principais fatores de perda pós-colheita e é determinado, na grande maioria das vezes, de maneira subjetiva, por meio da coloração externa e tamanho do fruto, não havendo uma padronização consensual entre os produtores. A utilização de índices de maturação permite que se conheça o ponto de colheita ideal para cada situação mercadológica.

A taxa de crescimento medida em frutos, com relação ao seu desenvolvimento, tem sido determinada com base no diâmetro, comprimento, circunferência, peso, volume, coloração e formato do fruto, desde a antese até a maturação (CASTRO NETO; REINHARDT, 2003). O potencial de conservação e a qualidade dos frutos são determinados, principalmente, pelo estádio de maturação em que os frutos são colhidos. Estudos sobre o desenvolvimento de frutos e sementes são importantes para estabelecer estratégias de colheita e técnicas adequadas de armazenamento pós–colheita para aumentar a vida útil e, com isso, melhorar o aproveitamento do potencial de comercialização do fruto (GURJÃO et al., 2006).

Segundo Loaiza-Echeverri (2011), o termo curva de crescimento evoca a imagem de uma curva sigmoide que descreve uma série de medidas de tamanho ao longo do tempo. Os trabalhos encontrados na área vegetal destacam que a análise de crescimento é fundamental, pois descreve as mudanças na produção vegetal em função do tempo (URCHEI; ROGRIGUES; STONE, 2000). Silva et al. (2011) citam que o crescimento da maioria das espécies pode ser descrito por uma curva sigmoide e, portanto, ajustado por modelos não lineares, também conhecidos como regressões não lineares. Esses modelos, quando ajustados aos dados de peso, altura, comprimento e outras características quantitativas de interesse, ao longo do tempo, tanto em animais quanto em vegetais, permitem sintetizar grande número de medidas e informações em apenas alguns parâmetros de interpretação biológica.

A esse respeito, a análise estatística é usada frequentemente para estudar os processos de crescimento envolvidos no sistema de produção vegetal. Existem muitas pesquisas que avaliam o crescimento e desenvolvimento de frutíferas, utilizando modelos estatísticos. Esses estudos têm demonstrado que o comportamento do crescimento e desenvolvimento de frutos podem, satisfatoriamente, ser descritos por modelos não lineares. Vale ressaltar que a maior parte desses trabalhos faz uso de modelos de regressão não linear univariados, aqueles que utilizam uma única característica que é avaliada em função do tempo (MARTINS FILHO et al., 2008; PRADO; SAVIAN; MUNIZ, 2013).

2.3 Modelos de regressão

Na estatística, a regressão é um método bastante utilizado que permite explorar e inferir a relação de uma (ou mais) variável resposta com variáveis preditoras específicas. As funções de regressões não lineares têm servido, e continuarão a servir, para descrever variáveis físicas e sistemas biológicos, além do mais, em tais modelos, os parâmetros têm interpretações práticas (GRAYBILL; IYER, 1994).

Em regressão, os modelos são classificados basicamente como: modelos lineares (aqueles que são lineares em relação aos parâmetros); modelos linearizáveis (aqueles que podem ser transformados em lineares, por meio de alguma transformação) e modelos não lineares (modelos em que pelo menos uma das derivadas parciais depende de algum parâmetro do modelo) (DRAPER; SMITH, 1998). Alguns modelos de regressão não linear são utilizados para representar curvas de crescimento e apresentam algumas peculiaridades quanto ao procedimento de ajuste.

De acordo com Seber e Wild (2003), uma vantagem dos modelos não lineares é que os seus parâmetros fornecem um maior conhecimento sobre o fenômeno em estudo, além de, geralmente, fornecerem um bom ajuste, com menos parâmetros do que os modelos lineares. Porém, uma das desvantagens de ajuste de modelos não lineares é que os ajustes do modelo, em geral, exige uma otimização iterativa de funções que, muitas vezes, não converge facilmente para a solução desejada. Os algoritmos de uso generalizado que resolvem esse problema são baseados em modificações e aproximações de Newton-Raphson e Gauss-Newton.

Sob o ponto de vista da inferência estatística clássica, para os modelos de regressão linear, os estimadores de mínimos quadrados apresentam propriedades ótimas, isto é, são não viesados, normalmente distribuídos e, ainda, possuem variância mínima entre qualquer outra classe de estimadores. Essas propriedades são consideradas como as melhores que uma classe de estimadores pode apresentar (SEARLE, 1971). Já, no caso de modelos não lineares, as propriedades somente são válidas assintoticamente, isto é, quando o tamanho da amostra é suficientemente grande. Em geral, em pequenas amostras essas propriedades não são satisfeitas.

Vários tipos de modelos podem ser usados, de acordo com suas características, para facilitar a interpretação dos processos envolvidos no sistema de produção vegetal. Esses modelos podem fornecer também informações e estimativas úteis, particularmente se forem baseados em parâmetros que permitem interpretação biológica. Em geral, uma característica comum aos modelos não lineares com três parâmetros, como uma das parametrizações do modelo logístico, é que eles mesmos utilizam dois parâmetros biológicos relevantes, e outro que identifica uma constante matemática. Outro fato é que embora as letras utilizadas para designar cada um dos parâmetros coincidam nos diferentes modelos, existem variações quanto às suas interpretações (FORNI, 2007). Os modelos não lineares empíricos usados frequentemente para estimar o crescimento de frutos incluem a função logística e de gompertz. Primeiramente, a função logística foi indicada para o estudo do crescimento de populações humanas por Verhulst em 1838. A função de gompertz foi estudada por

Gompertz, Winsor e Laird para descrever a taxa de mortalidade numa população (GOMPERTZ, 1825).

Na área vegetal vários autores têm trabalhado com modelos não lineares, como é o caso de Jesus, Martins e Barbosa (2008) que estudaram o desenvolvimento de frutos de Longan avaliaram várias características, físicas e químicas, e verificaram que a curva do desenvolvimento dos frutos segue um padrão sigmoidal simples, modelada por uma função composta envolvendo a logística e a exponencial quadrática.

Alvarez e Castro (1999) analisaram o comportamento do crescimento de cana crua e cana queimada nos primeiro e segundo anos de rebrota e verificaram que a altura das plantas apresentou curvas de crescimento em formato sigmoidal. Foi ajustado, dentre outros modelos, o modelo não linear de crescimento gompertz aos dados de matéria seca de folha e colmo, índice de área foliar e acúmulos de graus-dia para cana-de-açúcar.

Pereira et al. (2000) estudaram o comportamento da bananeira (*Musa* spp.) 'Prata Anã' no primeiro ciclo de produção sob diferentes espaçamentos e verificaram um padrão sigmoidal para o crescimento das plantas ajustando o modelo logístico para a característica altura das plantas. No estudo de crescimento de frutos de maracujá analisando as variáveis diâmetro, comprimento, volume e massa fresca em função dos dias, Alves et al. (2012) concluíram que o padrão de crescimento de todas as variáveis seguiu um padrão sigmoidal e utilizaram o modelo logístico na descrição das curvas de crescimento.

Fernandes et al. (2014) compararam o ajuste dos modelos de regressão não lineares logístico e gompertz, na descrição das curvas de crescimento do fruto do cafeeiro, considerando a heterogeneidade das variâncias amostrais, nas diferentes faces de exposição ao sol, sendo o gompertz ponderado, o melhor modelo para a descrição do crescimento do fruto. Terra, Muniz e Savian (2010) verificaram a qualidade do ajuste dos modelos logístico e gompertz, aos dados de crescimento, em comprimento e diâmetro dos frutos de tamareira-anã e destacaram que o modelo logístico se ajustou melhor aos dados.

Benassi et al. (2007) estudaram dentre outras características, o diâmetro externo longitudinal e transversal de frutos de coqueiro anão verde, ajustando o modelo logístico aos dados. Carlesso et al. (2007) compararam cinco modelos não lineares exponenciais e variações deste, com o objetivo de determinar o modelo que melhor descreve a secagem por convecção em camada fina de sementes de maracujá-amarelo.

Na literatura, poucos trabalhos envolvem estudo de modelos não lineares multivariados. A análise multivariada aborda um conjunto de métodos que permite a análise simultânea de duas ou mais variáveis para cada indivíduo ou objeto em análise e a seguir são apresentados alguns destes trabalhos.

Chiodi e Mineo (1997), estudando o crescimento de crianças, utilizaram o modelo não linear multivariado Von Bertalanffy, considerando uma função de autocorrelação num processo contínuo AR (1) para a série de resíduos. Os autores estimaram os parâmetros do modelo, por meio da maximização da função log-verossimilhança por meio dos dados experimentais peso e altura.

Xu et al. (2008) realizaram, pela primeira vez, uma análise não linear multivariada a dados longitudinais de neuroimagem, para estudar o desenvolvimento inicial do cérebro, utilizando o modelo logístico, abordando a metodologia de modelos mistos. Strathe et al. (2010) analisaram o particionamento de energia no crescimento de suínos por meio dos modelos não lineares mistos multivariado Michaelis Menten e Gompertz e a estimação dos parâmetros foi feita por meio do método de máxima versossimilhança.

2.4 Método de estimação por mínimos quadrados para modelo não linear multivariado e alguns aspectos teóricos

Gallant (1987) afirma que o método dos mínimos quadrados é utilizado na estimação dos parâmetros em modelos não lineares, da mesma maneira que em modelos lineares. Porém, para o sistema de equação normais não linear, não existe uma solução explícita a qual deve ser obtida utilizando de processos iterativos, usando um software estatístico.

Vários métodos iterativos são propostos na literatura para a obtenção das estimativas de mínimos quadrados dos parâmetros de um modelo de regressão não-linear. Os mais utilizados são o método de *Gauss-Newton* ou *método da linearização*, o método *Steepest-Descent* ou método *gradiente* e o método de *Marquardt* (BATES; WATTS, 1988).

Gallant (1987) utiliza dois modos para escrever o modelo não linear na forma de vetor. O primeiro enfatiza o fato de que o modelo consiste de *M* equações de regressão não linear univariadas separadas, representadas de modo geral por:

$$\underset{\sim}{y}_{\alpha} = \underset{\sim}{f}_{\alpha}(\underset{\sim}{x}_{t}, \underset{\sim}{\theta}_{\alpha}) + \underset{\sim}{e}_{\alpha}$$

Em que $\alpha = 1,\ldots,M$ representa cada característica avaliada na unidade experimental, t = 1, 2,..., n é o número de observações dentro de cada característica, $\underset{\sim}{x}_{t}$ é um vetor k-dimensional de variáveis regressoras, cada variável resposta separadamente $\underset{\sim}{y}_{\alpha}$ representa um vetor n-dimensional, cada $\underset{\sim}{\theta}_{\alpha}$ é um vetor de parâmetros desconhecidos p_{α}-dimensional e $\underset{\sim}{e}_{\alpha t}$ representa o vetor n-dimensional de erros experimentais de cada modelo.

Neste caso, a forma matricial deste modelo pode ser representada da seguinte forma:

$$\underset{\sim}{y}_{\alpha}=\begin{bmatrix} y_{\alpha 1} \\ y_{\alpha 2} \\ \vdots \\ y_{\alpha t} \\ \vdots \\ y_{\alpha n} \end{bmatrix}_{n\times 1} \quad \underset{\sim}{f}_{\alpha}(\underset{\sim}{x}_{t},\underset{\sim}{\theta}_{\alpha})=\begin{bmatrix} f_{\alpha}(\underset{\sim}{x}_{1},\underset{\sim}{\theta}_{\alpha}) \\ f_{\alpha}(\underset{\sim}{x}_{2},\underset{\sim}{\theta}_{\alpha}) \\ \vdots \\ f_{\alpha}(\underset{\sim}{x}_{t},\underset{\sim}{\theta}_{\alpha}) \\ \vdots \\ f_{\alpha}(\underset{\sim}{x}_{n},\underset{\sim}{\theta}_{\alpha}) \end{bmatrix}_{n\times 1} \quad \underset{\sim}{e}_{\alpha}=\begin{bmatrix} e_{\alpha 1} \\ e_{\alpha 2} \\ \vdots \\ e_{\alpha 2} \\ \vdots \\ e_{\alpha n} \end{bmatrix}_{n\times 1}$$

O segundo modelo de regressão não linear enfatiza a natureza multivariada dos dados, representado pela notação:

$$\underset{\sim}{y}_{t}=\underset{\sim}{f}(\underset{\sim}{x}_{t},\underset{\sim}{\theta})+\underset{\sim}{e}_{t}$$

Em que t =1,..,n, com $\underset{\sim}{y}_{t}$ um vetor M-dimensional e $\underset{\sim}{\theta}$ o vetor de parâmetros p-dimensional. Mais adiante, a notação multivariada $\underset{\sim}{y}_{t}$ será trabalhada em alguns aspectos com o objetivo de fazer um paralelo com algumas ideias aplicadas à notação do primeiro modelo de respostas univariadas $\underset{\sim}{y}_{\alpha}$.

No primeiro modelo de regressão não linear $\underset{\sim}{y}_{\alpha}$, a estimativa do vetor de parâmetros $\underset{\sim}{\theta}_{\alpha}$ de cada modelo é dada pelo vetor $\hat{\underset{\sim}{\theta}}_{\alpha}$ que minimiza a soma de quadrados do resíduo,

$$S\left(\underset{\sim}{\theta}_{\alpha}\right)=\left[\underset{\sim}{y}_{\alpha}-\underset{\sim}{f}_{\alpha}\left(\underset{\sim}{x}_{t},\underset{\sim}{\theta}_{\alpha}\right)\right]^{T}\left[\underset{\sim}{y}_{\alpha}-\underset{\sim}{f}_{\alpha}\left(\underset{\sim}{x}_{t},\underset{\sim}{\theta}_{\alpha}\right)\right]$$

Com isto, pode-se generalizar,

$$\underset{\sim}{\hat{e}}_\alpha = \underset{\sim}{y}_\alpha - \underset{\sim}{f}_\alpha\left(\underset{\sim}{x}_t, \underset{\sim}{\hat{\theta}}_\alpha\right) \qquad \alpha = 1, 2, \ldots, M$$

ou

$$\hat{E} = \left[\underset{\sim}{\hat{e}}_1\ \underset{\sim}{\hat{e}}_2 \cdots \underset{\sim}{\hat{e}}_M\right]$$

em que $\hat{E}$ é a matriz que representa as estimativas dos vetores de erros de cada modelo de regressão. Então,

$$\hat{\Sigma} = \frac{1}{n}\hat{E}^T\hat{E}$$

Em que $\hat{\Sigma}$ denota uma matriz de ordem M x M que representa a matriz de variância e covariância dos erros.

O conjunto de M regressões pode ser arranjado em uma única regressão $\underset{\sim}{y}$,

$$\underset{\sim}{y} = \underset{\sim}{f}\left(\underset{\sim}{\theta}\right) + \underset{\sim}{e}$$

Em que na forma matricial,

$$\underset{\sim}{y} = \begin{bmatrix} \underset{\sim}{y}_1 \\ \underset{\sim}{y}_2 \\ \vdots \\ \underset{\sim}{y}_M \end{bmatrix}_{nM\times 1} \quad \underset{\sim}{f}(\underset{\sim}{\theta}) = \begin{bmatrix} \underset{\sim}{f}_1(\underset{\sim}{\theta}_1) \\ \underset{\sim}{f}_2(\underset{\sim}{\theta}_2) \\ \vdots \\ \underset{\sim}{f}_M(\underset{\sim}{\theta}_M) \end{bmatrix}_{nM\times 1} \quad \underset{\sim}{e} = \begin{bmatrix} \underset{\sim}{e}_1 \\ \underset{\sim}{e}_2 \\ \vdots \\ \underset{\sim}{e}_M \end{bmatrix}_{nM\times 1} \quad \underset{\sim}{\theta} = \begin{bmatrix} \underset{\sim}{\theta}_1 \\ \underset{\sim}{\theta}_2 \\ \vdots \\ \underset{\sim}{\theta}_M \end{bmatrix}_{p\times 1}$$

com $p = \sum_{\alpha=1}^{M} p_{\alpha}$.

Nesta notação, a matriz de variância e covariância dos erros é dada por:

$$C\left(e,e^{T}\right)=\begin{pmatrix} \sigma_{11}I & \sigma_{12}I & \cdots & \sigma_{1M}I \\ \sigma_{21}I & \sigma_{22}I & \cdots & \sigma_{2M}I \\ \vdots & \vdots & & \vdots \\ \sigma_{M1}I & \sigma_{M2}I & \cdots & \sigma_{MM}I \end{pmatrix} = \sum \otimes I = \Omega$$

Em que $\otimes$ representa o operador produto de Kronecker; *I* é a matriz identidade de dimensões *n* x *n*; $\sum$ é a matriz de variâncias e covariâncias dos erros de dimensões *M* x *M*; *C* é a matriz resultante do produto de dimensões *nM* x *nM*.

Considerando a fatoração ou decomposição de Cholesky (FERREIRA, 2011) da inversa da matriz de covariâncias $\hat{\Sigma}^{-1} = \hat{\Lambda}^{T}\hat{\Lambda}$, em que $\hat{\Lambda}$ é o fator de Cholesky ($\hat{\Lambda}$ é uma matriz triangular superior com elementos diagonais estritamente positivos), o modelo rotacionado

$$\left(\hat{P}\otimes I\right)\underset{\sim}{y}=\left(\hat{P}\otimes I\right)\underset{\sim}{f}\left(\underset{\sim}{\theta}\right)+\left(\hat{P}\otimes I\right)\underset{\sim}{e}$$

pode ser reescrito por

$$\underset{\sim}{y}^{*} = \underset{\sim}{f}\left(\underset{\sim}{\theta}\right)^{*} + \underset{\sim}{e}^{*} .$$

Desse modo, pode-se mostrar que $C\left(\underset{\sim}{e}^{*},\underset{\sim}{e}^{*^{T}}\right) = {}_{nM}I_{nM}$ (ver GALLANT, 1987), ou seja, a matriz de variância e covariância entre os erros é uma matriz identidade e o modelo $\underset{\sim}{y}^{*} = \underset{\sim}{f}(\underset{\sim}{\theta})^{*} + \underset{\sim}{e}^{*}$ é um modelo não linear univariado em que $\underset{\sim}{\theta}$ pode ser estimado pela minimização

$$
\begin{aligned}
S(\underset{\sim}{\theta},\Sigma) &= \left[\underset{\sim}{y}^{*} - \underset{\sim}{f}(\underset{\sim}{\theta})^{*}\right]^{T}\left[\underset{\sim}{y}^{*} - \underset{\sim}{f}(\underset{\sim}{\theta})^{*}\right] = \\
&= \left[\underset{\sim}{y} - \underset{\sim}{f}(\underset{\sim}{\theta})\right]^{T}\left(\Sigma^{-1}\otimes I\right)\left[\underset{\sim}{y} - \underset{\sim}{f}(\underset{\sim}{\theta})\right]
\end{aligned}
$$

É claro que $\sum$ é desconhecido, então é feita a substituição de $\sum$ por $\hat{\Sigma}$ de forma que resta apenas estimar $\underset{\sim}{\theta}$ por $\underset{\sim}{\hat{\theta}}$ minimizando $S\left(\underset{\sim}{\theta},\hat{\Sigma}\right)$.

A decomposição de Cholesky da inversa da matriz de covariâncias é útil, pois faz com que o método de mínimos quadrados ordinários seja utilizado, ao invés do método de mínimos quadrados generalizados.

As ideias apresentadas acima podem ser implementadas ao adotar o esquema de notação multivariado. O modelo de regressão não linear multivariado $\underset{\sim}{y}_t = \underset{\sim}{f}(\underset{\sim}{x}_t,\underset{\sim}{\theta}) + \underset{\sim}{e}_t$, t = 1, 2,..., n, (GALLANT, 1987; SEBER; WILD, 2003) na forma matricial é

$$
\underset{\sim}{y}_t = \begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{Mt} \end{bmatrix}_{M\times 1}
\quad
\underset{\sim}{f}(\underset{\sim}{x}_t,\underset{\sim}{\theta}) = \begin{bmatrix} f_1(\underset{\sim}{x}_t,\underset{\sim}{\theta}_1) \\ f_2(\underset{\sim}{x}_t,\underset{\sim}{\theta}_2) \\ \vdots \\ f_M(\underset{\sim}{x}_t,\underset{\sim}{\theta}_M) \end{bmatrix}_{M\times 1}
\quad
\underset{\sim}{e}_t = \begin{bmatrix} e_{1t} \\ e_{2t} \\ \vdots \\ e_{Mt} \end{bmatrix}_{M\times 1}
\quad
\underset{\sim}{\theta} = \begin{bmatrix} \underset{\sim}{\theta}_1 \\ \underset{\sim}{\theta}_2 \\ \vdots \\ \underset{\sim}{\theta}_M \end{bmatrix}_{p\times 1}
$$

com t = 1, 2,..., n e $p = \sum_{\alpha=1}^{M} p_\alpha$.

Neste esquema de notação pode-se mostrar que:

$$S(\underset{\sim}{\theta},\Sigma)=\sum_{t=1}^{n}\left[\underset{\sim}{y}_t-\underset{\sim}{f}\left(\underset{\sim}{x}_t,\underset{\sim}{\theta}\right)\right]^T\left[\underset{\sim}{y}_t-\underset{\sim}{f}\left(\underset{\sim}{x}_t,\underset{\sim}{\theta}\right)\right]=$$

$$=\left[\underset{\sim}{y}-\underset{\sim}{f}\left(\underset{\sim}{\theta}\right)\right]^T\left(\Sigma^{-1}\otimes I\right)\left[\underset{\sim}{y}-\underset{\sim}{f}\left(\underset{\sim}{\theta}\right)\right]$$

Note que a soma de quadrados $S(\underset{\sim}{\theta},\Sigma)$ é o mesmo tanto para o modelo não linear univariado $\underset{\sim}{y}^*=\underset{\sim}{f}\left(\underset{\sim}{\theta}\right)^*+\underset{\sim}{e}^*$ quanto para o caso não linear multivariado $\underset{\sim}{y}_t=\underset{\sim}{f}(\underset{\sim}{x}_t,\underset{\sim}{\theta})+\underset{\sim}{e}_t$. Vale observar que o sistema de notação multivariada é mais conveniente quando as observações são retiradas do mesmo indivíduo, em conjunto, considerando um processo em série no tempo t = 1, 2,..., n, visto que as estimativas dos parâmetros serão mais precisas.

A soma de quadrados $S(\underset{\sim}{\theta},\Sigma)$ pode ser reescrita ou transformada de forma que permitirá o uso de programas de regressão não linear univariado para computação multivariada.

Observe que se $\hat{\Sigma}^{-1}$ é fatorado como $\hat{\Sigma}^{-1}=\hat{\Lambda}^T\hat{\Lambda}$ então,

$$S(\underset{\sim}{\theta},\Sigma)=\sum_{t=1}^{n}\left[\Lambda\underset{\sim}{y}_t-\Lambda\underset{\sim}{f}\left(\underset{\sim}{x}_t,\underset{\sim}{\theta}\right)\right]^T\left[\Lambda\underset{\sim}{y}_t-\Lambda\underset{\sim}{f}\left(\underset{\sim}{x}_t,\underset{\sim}{\theta}\right)\right]$$

Denotando $\Lambda_{(\alpha)}^T$ como α - ésima linha de Λ,

$$S(\underset{\sim}{\theta},\Sigma)=\sum_{t=1}^{n}\sum_{\alpha=1}^{M}\left[\Lambda_{(\alpha)}^T\underset{\sim}{y}_t-\Lambda_{(\alpha)}^T\underset{\sim}{f}\left(\underset{\sim}{x}_t,\underset{\sim}{\theta}\right)\right]^2$$

Agora, tem-se a soma de quadrados $S\left(\underset{\sim}{\theta},\Sigma\right)$ expressa como soma de quadrados de termos univariados e o que resta é encontrar um esquema de notação para remover a duplo somatório. Para esse fim, seja

$$
\begin{aligned}
& s = M\left(t-1\right)+\alpha \\
& \underset{\sim}{y}_s^* = \Lambda_{(\alpha)}^T \underset{\sim}{y}_t \\
& \underset{\sim}{x}_s^* = \left(\Lambda_{(\alpha)}^T, \underset{\sim}{x}_t^T\right)^T \\
& \underset{\sim}{f}_s^*\left(\underset{\sim}{x}_s^*, \underset{\sim}{\theta}\right) = \Lambda_{(\alpha)}^T \underset{\sim}{f}\left(\underset{\sim}{x}_t, \underset{\sim}{\theta}\right)
\end{aligned}
$$

para α =1,..., M e t = 1, 2,..., n tem-se

$$
S\left(\underset{\sim}{\theta},\Sigma\right)=\sum_{s=1}^{nM}\left[\underset{\sim}{y}_s^* - \underset{\sim}{f}_s^*\left(\underset{\sim}{x}_s^*, \underset{\sim}{\theta}\right)\right]^2
$$

Para a obtenção das estimativas dos parâmetros é necessário o uso de métodos iterativos, visto que, em geral, não se obtém uma solução explícita para o parâmetro por meio do método descrito acima para modelos não lineares.

2.5 Distribuição normal multivariada

A generalização da densidade normal univariada para duas ou mais dimensões desempenha um papel fundamental na análise multivariada, visto que a maioria das técnicas multivariadas parte do pressuposto de que os dados foram gerados de uma distribuição normal multivariada. Uma das principais razões para basear a análise estatística sobre a distribuição normal é que esse modelo probabilístico aproxima bem da distribuição de medições contínuas em muitas populações amostradas, ou seja, é a mais adequada para modelos populacionais em várias situações e a distribuição amostral de muitas estatísticas multivariadas é aproximadamente normal, independente da forma da distribuição da população original, em decorrência do efeito do limite central (ANDERSON, 1984; FERREIRA, 2011).

Sejam p variáveis normais independentes dadas por $Y_1, Y_2, ..., Y_p$ com densidade definida por:

$$f_{Y_i}(y_i) = 1 \Big/ \sqrt{2\pi\sigma_{ii}} \exp\left\{ -\frac{(y_i - \mu_i)^2}{2\sigma_{ii}} \right\}$$

Em que μ_i e σ_{ii} são a média e a variância, respectivamente, para a i-ésima variável. Ferreira (2011) define a densidade conjunta das p variáveis independentes dispostas no vetor aleatório $Y = \left[Y_1, Y_2, ..., Y_p \right]^T$ por:

$$f_Y(y) = \prod_{i=1}^{p} f(y_i) = (2\pi)^{-\frac{p}{2}} \left(\prod_{i=1}^{p} \sigma_{ii} \right)^{-\frac{1}{2}} \exp\left\{ -\sum_{i=1}^{p} \frac{(y_i - \mu_i)^2}{2\sigma_{ii}} \right\}$$

Arranjando a equação anterior, utilizando matrizes e vetores, obtém-se:

$$f_Y(y) = (2\pi)^{-\frac{p}{2}} \begin{vmatrix} \sigma_{11} & 0 & \dots & 0 \\ 0 & \sigma_{22} & \dots & 0 \\ \vdots & \dots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{pp} \end{vmatrix}^{-\frac{1}{2}} \times$$

$$\times \exp\left\{ -\frac{1}{2}(y-\mu)^T \begin{bmatrix} \sigma_{11} & 0 & \dots & 0 \\ 0 & \sigma_{22} & \dots & 0 \\ \vdots & \dots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_{pp} \end{bmatrix}^{-1} (y-\mu) \right\} =$$

$$= (2\pi)^{-\frac{p}{2}} |V|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2}\left((y-\mu)^T V^{-1} (y-\mu) \right) \right\}$$

Em que $V = diag(\sigma_{ii})$ e $\mu = [\mu_1, \mu_2, ..., \mu_p]^T$ são a matriz de covariâncias e o vetor de médias, respectivamente. Usando uma matriz de covariâncias positiva definida mais geral do que V dada por:

$$\Sigma = \begin{bmatrix} \sigma_{11} & \dots & \sigma_{1p} \\ \vdots & \ddots & \vdots \\ \sigma_{p1} & \dots & \sigma_{pp} \end{bmatrix}$$

Obtém-se a função densidade da distribuição normal multivariada (ANDERSON, 1984), dada por:

$$f_Y(y) = (2\pi)^{-\frac{p}{2}} |\Sigma|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2}\left((y-\mu)^T \Sigma^{-1} (y-\mu) \right) \right\}.$$

2.6 Distribuição normal bivariada

A distribuição normal bivariada (p = 2) é um caso particular da distribuição normal multivariada. Sejam as variáveis Y_1 e Y_2 com matriz de covariâncias:

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}$$

E vetor de médias $\mu = [\mu_1, \mu_2]^T$. Pode-se especificar ainda a covariância em função do coeficiente de correlação ρ_{12} (ou simplesmente ρ) por $\sigma_{12} = \rho\sqrt{\sigma_{11}\sigma_{12}}$ (FERREIRA, 2011), verificar que $|\Sigma| = \sigma_{11}\sigma_{12}\left(1-\rho^2\right)$ e que a inversa da matriz de covariâncias Σ^{-1} é:

$$\Sigma^{-1} = \frac{1}{1-\rho^2}\begin{bmatrix} \dfrac{1}{\sigma_{11}} & -\dfrac{\rho}{\sqrt{\sigma_{11}}\sqrt{\sigma_{12}}} \\ -\dfrac{\rho}{\sqrt{\sigma_{11}}\sqrt{\sigma_{12}}} & \dfrac{1}{\sigma_{22}} \end{bmatrix}$$

Além disso, a forma quadrática $(y-\mu)^T \Sigma^{-1}(y-\mu)$ pode ser expandida por:

$$(y-\mu)^T \Sigma^{-1}(y-\mu) =$$
$$= \frac{1}{1-\rho^2}\left[\left(\frac{y_1-\mu_1}{\sqrt{\sigma_{11}}}\right)^2 + \left(\frac{y_2-\mu_2}{\sqrt{\sigma_{22}}}\right)^2 - 2\rho\left(\frac{y_1-\mu_1}{\sqrt{\sigma_{11}}}\right)\left(\frac{y_2-\mu_2}{\sqrt{\sigma_{22}}}\right)\right]$$

Dessa forma, a função densidade da distribuição normal bivariada é dada por:

$$f\left(y_1, y_2\right)=\frac{1}{2\pi\sqrt{\sigma_{11}\sigma_{22}\left(1-\rho^2\right)}}\times$$
$$\times\exp\left\{-\frac{1}{2\left(1-\rho^2\right)}\left[\left(\frac{y_1-\mu_1}{\sqrt{\sigma_{11}}}\right)^2+\left(\frac{y_2-\mu_2}{\sqrt{\sigma_{22}}}\right)^2-2\rho\left(\frac{y_1-\mu_1}{\sqrt{\sigma_{11}}}\right)\left(\frac{y_2-\mu_2}{\sqrt{\sigma_{22}}}\right)\right]\right\}.$$

2.7 Avaliadores da qualidade de ajuste

A seleção de um modelo deve ser coerente com o evento biológico em estudo, além de basear-se em um adequado ajuste dos modelos aos dados amostrais. Os critérios de informação são utilizados nas mais diversas áreas das ciências e os autores Burnham e Anderson (2004) destacam a importância de selecionar modelos, baseando-se em princípios científicos.

Diversas são as abordagens utilizadas para esse fim tais como: critério de informação de Akaike (AIC), critério de Akaike corrigido (AIC_c) e critério de informação bayesiano (BIC) que, de acordo com Emiliano, Vivanco e Menezes (2014) esses critérios são os mais conhecidos e aplicados, sendo implementados na maioria dos softwares estatísticos. Segundo Burnham e Anderson (2004), o critério de informação de Akaike corrigido (AIC_c) é apenas uma correção do viés do AIC e a sua utilização é indicada quando a razão n/p é pequena, isto é, $n/p<40$. Além disso, quando essa relação é suficientemente grande, ambos os critérios, AIC e AIC_c, apresentam resultados semelhantes, e suas expressões são dadas respectivamente por:

$$AIC = \ln(QMR) + 2(p+1)/n$$

$$AIC_c = n\ln\left(\frac{SQR}{n}\right) + 2(p+1) + \frac{2(p+1)(p+2)}{n-(p+2)}$$

Em que: n é o número de observações (tamanho da amostra), QMR é o quadrado médio residual, SQR é a soma de quadrado residual e p é o número de parâmetros do modelo.

Para a escolha de modelos com diferentes números de parâmetros é indicado utilizar o coeficiente de determinação ajustado R^2_{aj}, pois ele leva em consideração a quantidade de parâmetros do modelo, diferentemente do coeficiente de determinação R^2 que não faz ponderações quanto ao número de parâmetros, e quanto maior o valor de R^2_{aj} melhor o ajuste do modelo (DRAPER; SMITH, 1998; RIBEIRO, 2015). Vale lembrar que o coeficiente de determinação R^2 expressa o quanto o modelo consegue explicar os valores observados e o coeficiente de determinação ajustado R^2_{aj} indica se um modelo proporcionou melhor ajuste do que outro modelo. O coeficiente de determinação ajustado é dado por:

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

Em que: n é o número de observações e p é o número de parâmetros do modelo.

Outro avaliador bastante utilizado para a seleção de modelo é o quadrado médio residual (QMR), calculado dividindo-se a soma de quadrados do resíduo, pelo número de observações, que é o estimador de máxima verossimilhança da variância residual (SOUZA et al., 2010). O desvio padrão residual DPR que é a raiz do quadrado médio residual, também muito utilizado, indica a distância entre as estimativas do modelo e as observações e quanto menor o valor do desvio padrão residual, melhor é o ajuste do modelo estimado (SOUZA et al., 2014).

3 MATERIAL E MÉTODOS

3.1 Dados experimentais

Os dados utilizados para o ajuste dos modelos foram obtidos de um experimento (BENASSI, 2006) em que foi mensurado, dentre outras características, os diâmetros externos, longitudinal (DEL) e transversal (DET), de frutos de coqueiro da variedade anã verde. O trabalho de campo foi realizado em um coqueiral (*Cocos nucifera* L.) variedade anã verde, com idade de sete anos, instalado na Estação Experimental de Citricultura de Bebedouro/SP – EECB. O espaçamento utilizado foi de 7,0m entre linhas e 6,0m entre plantas, em disposição triangular. Antes do plantio foi realizada a calagem em área total, objetivando-se elevar a saturação de bases para 65%. As covas foram abertas mecanicamente nas dimensões de 60 x 60 x 60 cm utilizando-se para seu preparo, 20 litros de esterco de curral curtido, 800 gramas de superfosfato simples e 300 gramas de calcário dolomítico.

As avaliações foram feitas em oito plantas diferentes, marcando-se uma inflorescência em cada planta, com base nas medidas de três frutos identificados nas regiões apical, mediana e basal do cacho, objetivando maior representatividade amostral, totalizando-se 24 frutos. A primeira medição foi feita no dia da abertura da inflorescência que se deu em 15/04/2004, e as demais, a cada 15 dias nos frutos marcados, encerrando-se em 01/05/2005, totalizando 26 medições. Em cada data foram anotados os diâmetros externos, longitudinal e transversal dos 24 frutos, obtendo-se as médias para o ajuste dos modelos.

3.2 Modelo logístico e gompertz univariado e multivariado

Aos dados, primeiramente, foram ajustados os modelos univariados logístico e gompertz considerando a estrutura autoregressiva de primeira ordem na série de resíduos.

1. Modelos univariados
   i. Modelo logístico-L

$$y_i = \frac{a}{\left\{1+b\left[\exp\left(-kt_i\right)\right]\right\}} + e_i$$

   ii. Modelo gompertz-G

$$y_i = a\exp\left\{-b\left[\exp\left(-kt_i\right)\right]\right\} + e_i$$

Em que i = 1, 2, 3,..., 26 $\left(t_1 = 1, t_2 = 15, t_3 = 30, ..., t_{26} = 375\right)$ dias após a abertura da inflorescência. Para os modelos acima são apresentadas as seguintes interpretações:

a) y_i expressa o valor da variável dependente (diâmetro externo longitudinal (DEL) e/ou diâmetro externo transversal (DET)) na i-ésima avaliação em cm;
b) a corresponde à assíntota do modelo, indicando o valor de estabilização da variável dependente em relação ao tempo, em cm;
c) b corresponde ao parâmetro de locação, sem interpretação biológica;

d) k refere-se à taxa de expansão do crescimento;

e) e_i é o erro aleatório associado à observação y_i .

Para a estrutura de erros tem-se que: $e_i = \varphi_1 e_{i-1} + ... + \varphi_q e_{i-q} + \varepsilon_i$, com i = 1,..., n e n o número de medições no fruto; e_i é o resíduo do ajuste na i-ésima medição; φ_1 é o parâmetro autorregressivo de ordem 1; e_{i-1} é o resíduo do ajuste no tempo imediatamente anterior a i-ésima medição; φ_q é o parâmetro autoregressivo de ordem q; e_{i-q} é o resíduo do ajuste em q tempos anterior a i-ésima medição; ε_i é o ruído branco com distribuição normal $N\left(0,\sigma^2\right)$. Para a estrutura de erros independentes os parâmetros autorregressivos serão nulos e, em consequência, $e_i = \varepsilon_i$ (MAZZINI et al., 2005).

Os modelos de crescimento não lineares bivariados, logístico, gompertz e híbridos (composição dos modelos logístico e gompertz combinados para as duas variáveis respostas), considerando estrutura de erros independentes e autorregressivas, foram ajustados. A seguir, são apresentados os componentes destes modelos citados.

2. Modelos multivariados

Os casos estudados foram:

i. <u>Modelo logístico para ambas as variáveis DEL e DET - LL</u>

ii. <u>Modelo gompertz para ambas as variáveis DEL e DET - GG</u>

iii. <u>Modelo logístico para o DEL e gompertz para o DET - LG</u>

iv. <u>Modelo gompertz para o DEL e logístico para o DET - GL</u>

Os modelos multivariados representados no geral por $\underset{\sim}{y}_t = \underset{\sim}{f}\left(\underset{\sim}{t},\underset{\sim}{\theta}\right)+\underset{\sim}{e}_t$ apresentam as seguintes componentes para cada modelo acima, dada da seguinte forma:

i.

$$\underset{\sim}{f}\left(\underset{\sim}{t},\underset{\sim}{\theta}\right)=\begin{pmatrix} \dfrac{a_1}{\left\{1+b_1\left[\exp\left(-k_1t\right)\right]\right\}} \\ \dfrac{a_2}{\left\{1+b_2\left[\exp\left(-k_2t\right)\right]\right\}} \end{pmatrix}$$

ii.

$$\underset{\sim}{f}\left(\underset{\sim}{t},\underset{\sim}{\theta}\right)=\begin{pmatrix} a_1\exp\left\{-b_1\left[\exp\left(-k_1t\right)\right]\right\} \\ a_2\exp\left\{-b_2\left[\exp\left(-k_2t\right)\right]\right\} \end{pmatrix}$$

iii.

$$\underset{\sim}{f}\left(\underset{\sim}{t},\underset{\sim}{\theta}\right)=\begin{pmatrix} \dfrac{a_1}{\left\{1+b_1\left[\exp\left(-k_1t\right)\right]\right\}} \\ a_2\exp\left\{-b_2\left[\exp\left(-k_2t\right)\right]\right\} \end{pmatrix}$$

iv.

$$\underset{\sim}{f}\left(\underset{\sim}{t},\underset{\sim}{\theta}\right)=\begin{pmatrix} a_1\exp\left\{-b_1\left[\exp\left(-k_1t\right)\right]\right\} \\ \dfrac{a_2}{\left\{1+b_2\left[\exp\left(-k_2t\right)\right]\right\}} \end{pmatrix}$$

Com vetor de parâmetros $\underset{\sim}{\theta}=\left(a_1,b_1,k_1,a_2,b_2,k_2\right)^T$, $\underset{\sim}{y}_t=\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix}$ as variáveis respostas DEL e DET respectivamente para y_{1t} e y_{2t} e $\underset{\sim}{e}_t=\begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix}$, em cada medição t = 1, 15, 30, ..., 375 dias. Para os modelos ii e iv foram consideradas estruturas de erros autorregressivas de segunda ordem, de modo que o vetor de parâmetros é dado por $\underset{\sim}{\theta}=(a_1,b_1,k_1,a_2,b_2,k_2,\varphi_1,\varphi_2)$ e a descrição da estrutura está apresentada nos modelos univariados. Vale ressaltar que a soma de quadrado do modelo multivariado foi descrita conforme o método apresentado anteriormente, ou seja, em função de termos não lineares univariados, de modo que a estrutura de erros autorregressiva foi considerada de maneira similar ao caso dos modelos univariados.

Com relação aos parâmetros, a_1 ou a_2, b_1 ou b_2, k_1 ou k_2 possuem as mesmas interpretações de a, b e k, dadas nos modelos univariados respectivos.

3.3 Método iterativo de Gauss-Newton

No presente estudo, foi utilizado o método de Gauss-Newton, descrito em Gallant (1987), que apresenta o seguinte desenvolvimento dos termos na notação $\underset{\sim}{f_s^*}\left(\underset{\sim}{x_s^*},\underset{\sim}{\theta}\right)$, dada, anteriormente, no método de mínimos quadrados:

$$D\left(\underset{\sim}{\theta},\Sigma\right)=\left[\sum_{s=1}^{nM}\left(\frac{\partial}{\partial\underset{\sim}{\theta}}\underset{\sim}{f_s^*}\left(\underset{\sim}{x_s^*},\underset{\sim}{\theta}\right)\right)\left(\frac{\partial}{\partial\underset{\sim}{\theta}}\underset{\sim}{f_s^*}\left(\underset{\sim}{x_s^*},\underset{\sim}{\theta}\right)\right)^T\right]^{-1}\times$$
$$\sum_{s=1}^{nM}\left(\frac{\partial}{\partial\underset{\sim}{\theta}}\underset{\sim}{f_s^*}\left(\underset{\sim}{x_s^*},\underset{\sim}{\theta}\right)\right)\left[\underset{\sim}{y_s^*}-\underset{\sim}{f_s^*}\left(\underset{\sim}{x_s^*},\underset{\sim}{\theta}\right)\right]=$$

$$=\left[\sum_{t=1}^{n}\sum_{\alpha=1}^{M}\left(\frac{\partial}{\partial \underset{\sim}{\theta}}\underset{\sim}{f}\left(\underset{\sim}{x}_t,\underset{\sim}{\theta}\right)\right)\Lambda_{(\alpha)}\Lambda_{(\alpha)}^{T}\left(\frac{\partial}{\partial \underset{\sim}{\theta}}\underset{\sim}{f}\left(\underset{\sim}{x}_t,\underset{\sim}{\theta}\right)\right)\right]^{-1}\times$$

$$\sum_{t=1}^{n}\sum_{\alpha=1}^{M}\left(\frac{\partial}{\partial \underset{\sim}{\theta}}\underset{\sim}{f}\left(\underset{\sim}{x}_t,\underset{\sim}{\theta}\right)\right)\Lambda_{(\alpha)}\Lambda_{(\alpha)}^{T}\left[\underset{\sim}{y}_t-\underset{\sim}{f}\left(\underset{\sim}{x}_t,\underset{\sim}{\theta}\right)\right]=$$

$$=\left[\sum_{t=1}^{n}\left(\frac{\partial}{\partial \underset{\sim}{\theta}}\underset{\sim}{f}\left(\underset{\sim}{x}_t,\underset{\sim}{\theta}\right)\right)\Sigma^{-1}\left(\frac{\partial}{\partial \underset{\sim}{\theta}}\underset{\sim}{f}\left(\underset{\sim}{x}_t,\underset{\sim}{\theta}\right)\right)\right]^{-1}\times$$

$$\sum_{t=1}^{n}\left(\frac{\partial}{\partial \underset{\sim}{\theta}}\underset{\sim}{f}\left(\underset{\sim}{x}_t,\underset{\sim}{\theta}\right)\right)\Sigma^{-1}\left[\underset{\sim}{y}_t-\underset{\sim}{f}\left(\underset{\sim}{x}_t,\underset{\sim}{\theta}\right)\right]$$

Com o vetor de correção acima, o algoritmo de Gauss-Newton modificado para minimizar $S\left(\underset{\sim}{\theta},\Sigma\right)$ é:

a) Escolha o valor arbitrário inicial $\underset{\sim}{\theta}_0$. Compute $D_0 = D\left(\underset{\sim}{\theta}_0,\Sigma\right)$ e encontre $0<\lambda_0<1$ tal que $S\left(\underset{\sim}{\theta}_0+\lambda_0 D_0,\Sigma\right)<S\left(\underset{\sim}{\theta}_0,\Sigma\right)$;

b) Seja $\underset{\sim}{\theta}_1=\underset{\sim}{\theta}_0+\lambda_1 D_0$. Compute $D_1 = D\left(\underset{\sim}{\theta}_1,\Sigma\right)$ e encontre $0<\lambda_1<1$ tal que $S\left(\underset{\sim}{\theta}_1+\lambda_1 D_1,\Sigma\right)<S\left(\underset{\sim}{\theta}_1,\Sigma\right)$;

c) Seja $\underset{\sim}{\theta}_2=\underset{\sim}{\theta}_1+\lambda_2 D_1$. ...

Este processo é repetido até que algum critério de convergência seja aceito, isto é, $\underset{\sim}{\theta}_m \rightarrow \hat{\underset{\sim}{\theta}}$ em que *m* é o número de vezes que o processo foi repetido e $\hat{\underset{\sim}{\theta}}$ é o vetor de parâmetros estimado obtido. No SAS, o critério de convergência usado foi de 10^{-8}.

3.4 Procedimentos de ajuste

O teste de Durbin-Watson (DW) (SEBER; WILD, 2003) foi utilizado para avaliar a existência de autocorrelação residual. A estatística do teste é dada por:

$$DW = \frac{\sum_{i=2}^{n}\left(e_i - e_{i-1}\right)^2}{\sum_{i=2}^{n} e_i^2}$$

Em que e_i é o resíduo no tempo *i* e e_{i-1} é o resíduo no tempo *i-1*. Dessa forma, as estimativas dos parâmetros dos modelos foram obtidas considerando a estrutura de autocorrelação residual. O numerador representa a diferença ao quadrado entre dois resíduos sucessivos, somados desde a segunda observação até a *n*-ésima observação. O denominador representa a soma dos quadrados dos resíduos. Quando resíduos sucessivos são positivamente autocorrelacionados, o valor de DW irá se aproximar de 0. Se os resíduos não forem correlacionados, o valor de DW estará próximo de 2 (SUBRAMANIAM; SILVA; COUTINHO, 2007).

A verificação da presença de autocorrelação residual foi feita da seguinte forma: primeiramente, realizou-se o ajuste do modelo considerando a estrutura de erros independentes. Para esse vetor de resíduos foi verificada a significância da estatística do teste de Durbin-Watson (DW), que testa a hipótese nula de que os resíduos são não correlacionados contra a hipótese alternativa de que os resíduos seguem um processo autorregressivo de primeira ordem (AR1). Sendo a estatística do teste significativa, considerando um nível de significância de 5%, o vetor de resíduos foi utilizado no banco de dados como uma covariável, defasada em uma observação e indexada pelo parâmetro φ_1, e os parâmetros do modelo foram estimados conjuntamente. Ao realizar o teste de Durbin-Watson considerado no modelo, sendo a estatística do teste significativa, o novo vetor de resíduos é utilizado como uma covariável, defasada em duas observações e, assim, o parâmetro φ_2 é incluído no modelo.

No presente estudo, a seleção do modelo multivariado mais adequada para explicar o crescimento do fruto teve como base os seguintes avaliadores:

i-Critério de informação de Akaike corrigido (AICc) dado por meio da expressão (MOTULSKY; CHRISTOPOULOS, 2003):

$$AIC_c = n\ln\left(\frac{SQR}{n}\right) + 2\left(p+1\right) + \frac{2\left(p+1\right)\left(p+2\right)}{n-\left(p+2\right)}$$

Em que: n é o número de observações, SQR é a soma de quadrado residual e p é o número de parâmetros do modelo;

ii-Coeficiente de determinação ajustado R_{aj}^2, dado por:

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

Em que R^2 é o coeficiente de determinação não ajustado, n é o número de medições no fruto e p é o número de parâmetros do modelo;

iii-Quadrado médio residual (QMR), calculado, dividindo-se a soma de quadrados do resíduo pelo grau de liberdade do resíduo.

Foram construídos intervalos de 95% de confiança para os parâmetros do modelo multivariado considerado o mais adequado. De acordo com Gallant (1987) a situação mais comum é quando se deseja definir um intervalo de confiança em um dos componentes $\hat{\theta}_i$ do vetor $\underset{\sim}{\hat{\theta}}$. Dessa forma, o intervalo de confiança pode ser definido por:

$$\hat{\theta}_i \pm t_{\alpha/2}\sqrt{s^2\hat{c}_{ii}}$$

Em que $\hat{c}_{ii}$ é o *i*-ésimo elemento da diagonal de $\hat{C} = \left[F^T\left(\underset{\sim}{\hat{\theta}}\right)F\left(\underset{\sim}{\hat{\theta}}\right)\right]^{-1}$ e $F\left(\underset{\sim}{\hat{\theta}}\right)$ é a matriz de derivadas parciais do modelo, s^2 é o quadrado médio residual . Pode-se calcular $s^2\hat{c}_{ii}$ pela fórmula $s^2\hat{c}_{ii} = \hat{\sigma}_i\hat{\sigma}_j\rho_{ij}$, ou seja, o produto do erro padrão assintótico da estimativa do *i*-ésimo com o *j*-ésimo parâmetro e a correlação entre eles.

Considerando os dois diâmetros, DEL e DET, utilizados no estudo, e assumindo normalidade multivariada dos resíduos é feita toda a inferência, teste de hipóteses e intervalos de confiança.

Os procedimentos de ajuste dos modelos de crescimento mencionados foram implementados, por meio do programa computacional Statistical Analysis System SAS® (STATISTICAL ANALYSIS SYSTEM INSTITUTE, 2014), cuja rotina está no Anexo A.

4 RESULTADOS E DISCUSSÃO

4.1 Matriz de covariância dos erros e ajuste dos modelos não lineares univariados

Para execução do ajuste dos modelos não lineares multivariados foi necessário obter o vetor de resíduos dos modelos não lineares univariados para cada diâmetro externo com o intuito de obter a matriz de covariância dos erros. No caso dos ajustes individuais, foi considerado um processo autorregressivo de primeira ordem para os resíduos sendo que a escolha da ordem foi definida de acordo com a significância das estimativas dos parâmetros autoregressivos. A partir das matrizes de covariâncias dos erros obteve-se a matriz $\hat{P}$ (fator de Cholesky) para cada caso, objetivando a elaboração das rotinas para o ajuste dos modelos multivariados.

A estimativa da matriz de covariância dos erros $\hat{\Sigma}$ e do fator de Cholesky $\hat{P}$, para cada caso, considerando os diâmetros é dada a seguir:

i- Modelo logístico para ambas as variáveis DEL e DET - LL:

$$\hat{\Sigma} = \begin{bmatrix} 0,1183066 & 0,0731555 \\ 0,0731555 & 0,1668474 \end{bmatrix} \qquad \hat{\Lambda} = \begin{bmatrix} 3,4054014 & -1,493124 \\ 0 & 2,4481629 \end{bmatrix}$$

ii- Modelo gompertz para ambas as variáveis DEL e DET - GG:

$$\hat{\Sigma} = \begin{bmatrix} 0,1700826 & 0,1358164 \\ 0,1358164 & 0,2095969 \end{bmatrix} \qquad \hat{\Lambda} = \begin{bmatrix} 3,4905518 & -2,261837 \\ 0 & 2,1842764 \end{bmatrix}$$

iii- Modelo logístico para a variável DEL e gompertz para a variável DET - LG:

$$\hat{\Sigma} = \begin{bmatrix} 0,1183066 & 0,0706926 \\ 0,0706926 & 0,2095969 \end{bmatrix} \qquad \hat{\Lambda} = \begin{bmatrix} 3,2536288 & -1,09738 \\ 0 & 2,1842764 \end{bmatrix}$$

iv- Modelo gompertz para a variável DEL e logístico para a variável DET - GL:

$$\hat{\Sigma} = \begin{bmatrix} 0,1700826 & 0,1224286 \\ 0,1224286 & 0,1668474 \end{bmatrix} \qquad \hat{\Lambda} = \begin{bmatrix} 3,5300782 & -2,590286 \\ 0 & 2,4481629 \end{bmatrix}$$

Nos ajustes individuais foi necessário um processo autoregressivo de primeira ordem para os resíduos e na tabela 1 são apresentadas as estimativas dos parâmetros para os modelos univariados de acordo com cada característica:

Tabela 1 Estimativas dos parâmetros (erro padrão assintótico) dos modelos não lineares univariados logístico e gompertz aos dados de diâmetro, longitudinal (DEL) e transversal (DET), considerando a estrutura autoregressiva de primeira ordem para os resíduos

Estimativas	DEL		DET	
	Logístico	Gompertz	Logístico	Gompertz
$\hat{a}$	21,43	24,33	15,81	18,37
	(1,4920)	(2,4938)	(0,6218)	(9,3833)
$\hat{b}$	11,07	2,70	8,33	2,25
	(2,0931)	(0,2344)	(1,4049)	(1,63)
$\hat{k}$	0,0147	0,0075	0,0131	0,0063
	(0,00226)	(0,00119)	(0,000996)	(0,00177)
$\hat{\varphi}$	0,64	0,85	0,88	0,92
	(0,1382)	(0,1428)	(0,1764)	(0,6959)

Como objetivou-se, neste trabalho o ajuste dos modelos não lineares multivariados, a discussão será feita com base na tabela 2 a seguir.

4.2 Ajuste dos modelos não lineares multivariados

No ajuste dos modelos multivariados e estimação dos parâmetros, utilizou-se o PROC MODEL do programa SAS e a seleção do modelo mais adequado teve como base a precisão dos ajustes obedecendo aos critérios adotados. Em Gallant (1987) são apresentadas rotinas do SAS aplicados em um exemplo particular, para a estimação dos parâmetros de modelos multivariados.

Para os modelos GG e GL, considerando estrutura de erros independentes, o valor da estatística DW encontrava-se no valor de 2,77 e 2,96, respectivamente, indicando a presença de autocorrelação residual e, consequentemente, a necessidade de incorporar um processo autoregressivo AR para a série de resíduos. Ambos os valores são significativos (p-valor <0,0001) para o teste de autocorrelação positiva, indicando a necessidade de considerar um processo autorregressivo (AR), em princípio de primeira ordem, para a série de resíduos. Foi necessário o ajuste de um modelo AR2 para que as novas estatísticas do teste DW (2,23 para o modelo GG e 2,31 para o modelo GL) fossem não significativas (p-valor=0,8777 para o modelo GG e p-valor=0,0759 para o modelo GL).

Para o ajuste dos modelos LL e LG considerando a estrutura de erros independentes, o valor da estatística DW foi de aproximadamente 2,10 para o modelo LL e 1,76 para o modelo LG. Ambos os valores são não significativos (p-valor=0,6998 para LL e p-valor=0,2223 para LG) para o teste de autocorrelação positiva, indicando a não necessidade de considerar um processo autorregressivo (AR) na série de resíduos.

De acordo com os resultados apresentados na tabela 2, observa-se que o ajuste de todos os modelos multivariados foi bem satisfatório no estudo das características do crescimento dos frutos de coqueiro.

Avaliando os valores apresentados (Tabela 2), pode-se notar que as estimativas dos parâmetros $\hat{a}_1$ e $\hat{a}_2$ do modelo multivariado GG sofreram um decréscimo de 13% e 16%, aproximadamente, em relação às estimativas dos parâmetros *a* para o modelo gompertz univariado, em cada variável DEL e DET (Tabela 1). Já, no caso do modelo multivariado logístico LL as estimativas sofreram um acréscimo, com relação aos mesmos parâmetros, de 0,5% para $\hat{a}_1$ e 0,4% para $\hat{a}_2$ aproximadamente.

Tabela 2 Estimativas dos parâmetros (erro padrão assintótico) e avaliadores dos modelos multivariados logístico, gompertz e híbridos aos dados de diâmetro externo longitudinal (DEL) e transversal (DET) de frutos de coqueiro anão verde considerando a estrutura de erros autorregressivos AR2 apenas em GG e GL

Estimativas/ Avaliadores	MODELOS			
	LL	GG	LG	GL
$\hat{a}_1$	21,54	21,54	21,55	21,58
	(0,4269)	(0,6870)	(0,3890)	(0,6811)
$\hat{b}_1$	11,22	10,65	11,21	10,59
	(0,9305)	(1,4044)	(0,8451)	(1,3709)
$\hat{k}_1$	0,0147	0,0144	0,0146	0,0143
	(0,00068)	(0,00104)	(0,000617)	(0,00102)
$\hat{a}_2$	15,87	15,90	15,88	15,93
	(0,5124)	(1,2770)	(0,5232)	(1,2775)
$\hat{b}_2$	11,76	8,39	11,75	8,20
	(1,6004)	(1,9545)	(1,6290)	(1,8168)
$\hat{k}_2$	0,0147	0,0131	0,014717	0,0130
	(0,00110)	(0,00199)	(0,00112)	(0,00195)
$\hat{\varphi}_1$	---	-0,13	---	-0,16
		(0,1114)		(0,1134)
$\hat{\varphi}_2$	---	0,75	---	0,74
		(0,1111)		(0,1135)
R^2_{aj}	0,9867	0,9921	0,9898	0,9909
DW	2,1059	2,2265	1,7600	2,3128
QMR	2,5860	1,0390	2,1412	1,1846
AIC_c	59,55	15,59	49,76	22,41

Para o gompertz GG, os parâmetros de locação $\hat{b}_1$ e $\hat{b}_2$ apresentaram estimativas maiores (Tabela 2) se comparados com os resultados apresentados na tabela 1, para o modelo gompertz univariado. Esse parâmetro é estabelecido de acordo com os valores iniciais da variável resposta e o tempo e é importante para modelar a curva sigmoidal desde o início até a fase final (fruto totalmente desenvolvido) (SOUZA et al., 2010). No caso do modelo logístico LL, o parâmetro $\hat{b}_1$ é próximo da estimativa do parâmetro $\hat{b}$ do modelo logístico univariado e para $\hat{b}_2$ a estimativa diferiu bastante da estimativa obtida para o caso univariado, cuja diferença foi de 3,43, aproximadamente (Tabela 1).

Os parâmetros $\hat{k}_1$ e $\hat{k}_2$ representam a taxa de expansão do crescimento e indicam a velocidade de crescimento para atingir o tamanho assintótico de ambos os diâmetros. Na estimativa do parâmetro $\hat{k}_1$, observou-se que o modelo LL (0,0147) foi superior aos modelos GG (0,0144), GL (0,0146) e LG (0,0143). Já para $\hat{k}_2$ o modelo GL apresentou maior estimativa, que foi de 0,014717 (Tabela 2). Quanto maior o valor desse parâmetro, maior a velocidade de crescimento do fruto. Benassi (2007), estudando o diâmetro externo longitudinal de frutos de coqueiro anão verde obteve uma estimativa próxima ao do presente estudo, que foi de 0,01462 para a taxa de crescimento *k*, considerando o modelo logístico.

A componente autorregressiva de segunda ordem φ_2 estimada em ambos os modelos, GG e GL, no valor de 0,75 para o modelo GG e 0,74 para o modelo GL, indica uma alta correlação de segunda ordem entre os resíduos, confirmando a necessidade em ser considerado no ajuste do modelo. Pela tabela 3, por meio do intervalo de confiança, também se pode confirmar a conveniência do uso desta componente.

Para a maioria dos parâmetros dos modelos GG e LL, as estimativas dos erros padrões assintóticos dos mesmos, sofreram reduções bem expressivas ao ajustar os modelos multivariados, observando os valores dos modelos univariados. Em particular, para o parâmetro $\hat{a}_2$ do modelo multivariado gompertz verificou-se uma redução de aproximadamente 86% na estimativa do erro padrão assintótico, quando comparado ao erro padrão do parâmetro *a* do modelo gompertz univariado, para a característica DET (Tabela 2).

Ao avaliar os modelos multivariados logístico, gompertz e híbridos, os resultados da tabela 2 indicam que o modelo gompertz GG é considerado o mais adequado para descrever o crescimento de frutos de coqueiro da variedade anão verde quanto aos diâmetros externos longitudinais e transversais, pois apresentou o menor valor para AIC_c e QMR como também maior valor para o coeficiente de determinação ajustado R_{aj}^2.

É interessante atentar que os avaliadores AIC e R_{aj}^2 baseiam-se na sua capacidade de considerar o número de parâmetros dos modelos (DETMANN et al., 2012; RIBEIRO, 2015) e estes critérios, incluindo o QMR, são utilizados como critério se seleção de modelos com diferentes quantidades de parâmetros (SOUZA et al., 2010).

O intervalo de confiança para os parâmetros do modelo multivariado gompertz considerando os diâmetros externos longitudinais e transversais são apresentados na tabela 3 a seguir:

Tabela 3 Intervalo de 95 % confiança para os parâmetros (erro padrão assintótico) do modelo multivariado gompertz em que LI e LS é o limite inferior e superior, respectivamente

Parâmetros	LI	Estimativas	LS
$\hat{a}_1$	20,1904	21,54 (0,6870)	22,8836
$\hat{b}_1$	7,8967	10,65 (1,4044)	13,4020
$\hat{k}_1$	0,0123	0,0144(0,00104)	0,0164
$\hat{a}_2$	13,3925	15,90 (1,2770)	18,3982
$\hat{b}_2$	4,5603	8,39 (1,9545)	12,2219
$\hat{k}_2$	0,00916	0,0131(0,00199)	0,0169
$\hat{\varphi}_1$	-0,3475	-0,13 (0,1114)	0,0890
$\hat{\varphi}_2$	0,5354	0,75(0,1111)	0,9711

Com relação às estimativas dos parâmetros do modelo gompertz GG, verificou-se que os frutos apresentaram diâmetro externo longitudinal em torno de 21,54 cm e diâmetro externo transversal em torno de 15,90 cm. No decorrer do desenvolvimento, o fruto de coco inicialmente cresce mais em comprimento do que em largura, ou seja, com o diâmetro transversal menor que o diâmetro longitudinal e os frutos devem ser colhidos com o máximo de cuidado para evitar injúrias mecânicas provocadas pela queda. (MACIEL et al., 2009; PRADO; SAVIAN; MUNIZ, 2013).

O ponto de inflexão, que é o ponto em que a taxa de crescimento instantânea passa de crescente para decrescente, para o modelo gompertz GG se deu no valor de 9 cm e 5,8 cm em 71 dias (2,4 meses) para o diâmetro longitudinal e transversal respectivamente.

Prado, Savian e Muniz (2013), estudando o crescimento de frutos de coqueiro da variedade anã verde obtiveram para o modelo gompertz univariado, valores de 23,71 cm e 17,50 cm para os diâmetros externos longitudinais e transversais, respectivamente, e os autores concluíram que o modelo logístico univariado foi considerado o melhor modelo para a descrição das características. Vale observar que, no presente estudo, as estimativas foram menores (Tabela 2) considerando o modelo não linear multivariado gompertz e, além disso, este modelo foi considerado o mais adequado à descrição dos dados, discordando do estudo citado para o caso univariado, sobre o melhor modelo.

As estimativas obtidas com base no modelo ajustado estão em conformidade com a descrição das características das variedades de coqueiros, na literatura científica. Silva et al. (2013), avaliando características físicas de frutos de coqueiro anão verde obtiveram valores médios de 15,41 cm e 19,84 cm para o diâmetros transversal e longitudinal, respectivamente. Prado et al. (2013), estudando o crescimento de frutos da mesma variedade, obteve valor aproximado de 20,04 cm para o diâmetro longitudinal, em torno de 300 dias (10 meses), após a abertura da inflorescência. Silva et al. (2009) estudando características biométricas de frutos de coqueiro anão verde, com sete meses de idade (210 dias), obteve um valor de 17,23 cm para o diâmetro longitudinal e 14,97 cm para o diâmetro transversal, no sistema de produção convencional. Com base nestas informações, pode-se concluir que os resultados obtidos no presente estudo, de 21,54 cm para o DEL e 15,90 cm para o DET, concordam com as estimativas obtidas nos trabalhos encontrados na literatura científica.

Segundo Maciel (2008), os alimentos estão em constante atividade biológica, quer sejam industrializados ou não, o que se manifesta por alterações na natureza química, física e na atividade microbiológica ou enzimática. Tais alterações prejudicam sua qualidade, principalmente no caso dos alimentos de origem vegetal que, mesmo quando retirados das plantas, continuam respirando e, assim, realizando o metabolismo. Isso pode se caracterizar pela infestação por microorganismos, insetos ou pela presença de certos contaminantes químicos, os quais causam a perda de certos atributos específicos, tais como a cor, sabor, textura e viscosidade.

A colheita dos frutos em estádios adequados de maturação é determinante para a manutenção da qualidade pós-colheita. Frutos colhidos imaturos, além da baixa qualidade sensorial, são suscetíveis à desidratação e desordens fisiológicas (DREHMER; AMARANTE, 2008).

De acordo com Prado, Savian, Muniz (2013) os pesquisadores que estudam as características biométricas de frutos, conduzem os experimentos de avaliação até por volta de 210-240 dias após a abertura de inflorescência e o que se verifica é que os frutos mantidos na planta continuam a completar seu crescimento até períodos mais tardios. Portanto, deve-se ter cautela ao sugerir o ponto ideal de colheita dos frutos apenas com base nas medidas do fruto à maturação.

Aroucha et al. (2014), estudando análise físico-química e sensorial de água de coco das cultivares de coco anão verde e vermelho, concluíram que a cultivar anão verde e o 8° mês de maturação do fruto apresentou maior teor de frutose e melhor análise sensorial da água de coco que a cultivar anão vermelho, em que se obteve ótima aceitação sendo a mais preferida.

Maciel et al. (2009), avaliando a caracterização dos frutos de seis cultivares anão em diferentes estágios de desenvolvimento, afirmam que a melhor época para a colheita dos frutos foi no período de 168 aos 189 dias após a abertura da inflorescência, em que os mesmos atingiram os melhores atributos físicos. Benassi (2007), estudando frutos de coqueiro anão verde recomenda a colheita aos 8,5 meses de idade, considerando-se as propriedades desejáveis como o sabor e o volume da água de coco. De acordo com Resende et al. (2002), quando a produção se destina ao consumo *in natura* da água de coco, este deve ser colhido entre os seis e oito meses de idade, o que, em geral, é um consenso entre os pesquisadores desse fruto.

5 CONCLUSÃO

De acordo com os resultados obtidos no presente estudo, pode-se concluir que:

a) Apenas os modelos não lineares multivariados gompertz GG e híbrido GL apresentaram a necessidade de se incorporar um processo autorregressivo na série de resíduos, no caso, de segunda ordem;
b) Todos os modelos multivariados ajustados, logístico LL, gompertz GG e híbridos LG/GL, apresentaram um bom ajuste para a descrição da curva de crescimento dos frutos, apresentando adequada interpretação biológica dos parâmetros;
c) O modelo multivariado gompertz GG, considerando estrutura autorregressiva de segunda ordem, foi o mais adequado para ajustar os dados experimentais de diâmetro externo longitudinal (DEL) e transversal (DET), resultando em estimativas dos parâmetros bastante coerentes com os apresentados na literatura.

REFERÊNCIAS

ALVAREZ, I. A.; CASTRO, P. R. de C. Crescimento da parte aérea de cana crua e queimada. **Scientia Agrícola**, Piracicaba, v. 56, n. 4, p. 1069-1079, out./dez. 1999.

ALVES, R. R. et al. Desenvolvimento do maracujá doce em Viçosa, Minas Gerais. **Revista Ceres**, Viçosa, v. 59, n. 6, p. 127-133, nov./dez. 2012.

ANDERSON, T. W. **An introduction to multivariate statistical analysis**. 2. ed. New York: John Wiley & Sons, 1984. 675 p.

ARAGÃO, W. M. Cultivares de coqueiro. In: FONTES, H. R.; FERREIRA, J. M. S.; SIQUEIRA, L. A. (Ed.). **A cultura do coqueiro.** Aracaju: Embrapa Tabuleiros Costeiros, 2007. (Embrapa Tabuleiros Costeiros. Sistemas de Produção, 1). Disponível em: <http://sistemasdeproducao.cnptia.embrapa.br/FontesHTML/Coco/AculturadoCoqueiro/cultivares.htm>. Acesso em: 15 out. 2015.

AROUCHA, E. M. M. et al. Análise físico-química e sensorial de água-de-coco em função de estádio de maturação das cultivares de coco anão verde e vermelho. **Revista Agropecuária Científica no Semi-Árido**, Campina Grande, v. 10, n. 1, p. 33-38, jan./mar. 2014.

BATES, D. M.; WATTS, D. G. **Nonlinear regression analysis and its applications**. New York: John Wiley & Sons, 1988. 365 p.

BENASSI, A. C. **Caracterizações biométricas, química e sensorial de frutos de coqueiro variedade anã verde**. 2006. 49 p. Tese (Doutorado em Produção Vegetal) – Universidade Estadual Paulista "Júlio de Mesquita Filho", Jaboticabal, 2006.

BENASSI, A. C. et al. Caracterização biométrica de frutos de coqueiro, Vocos nucifera L. variedade anã-verde, em diferentes estádios de desenvolvimento. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 29, n. 2, p. 302-307, ago. 2007.

BITENCOURT, D. V.; PEDROTI, A. Usos da casca de coco: estudo das viabilidades de implantação de usina de beneficiamento de fibra de coco em Sergipe. **Revista da FAPESE de Pesquisa e Extensão**, Sergipe, v. 4, p. 115-124, jul./dez. 2008.

BURNHAM, K. P.; ANDERSON, D. R. Multimodel inference: understanding aic and bic in model selection. **Sociological Methods and Research**, Beverly Hills, v. 33, n. 2, p. 261-304, nov. 2004.

CARDOSO, G. D. et al. Uso da análise de crescimento não destrutiva como ferramenta para avaliação de cultivares. **Revista de Biologia e Ciências da Terra**, Campina Grande, v. 6, n. 2, p. 79-84, jul./dez. 2006.

CARLESSO, V. de O. et al. Avaliação de modelos de secagem em camada fina de sementes de maracujá amarelo. **Revista Brasileira de Sementes**, Londrina, v. 29, n. 2, p. 28-37, ago. 2007.

CASTRO NETO, M. T. de.; REINHARDT, D. H. Relações entre parâmetros de crescimento do fruto da manga cv. Haden. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 25, n. 1, p. 36-38, abr. 2003.

CAVALINI, F. C. et al. Maturity indexes for ‘Kumagai’ and ‘Paluma’ guavas. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 28, p. 176-179, ago. 2006.

CHIODI, M.; MINEO, A. M. The analysis of auxological data by means of nonlinear multivariate growth curves. In: VICHI, M.; OPITZ, O. **Classification and data analysis:** theory and application proceedings of the biannual meeting of the classification group of Società Italiana di Statistica (SIS) Pescara. Berlin: Springer-Verlag Berlin, 1997, p. 247-254.

DETMANN, E. et al. Uso de técnicas de regressão na avaliação, em bovinos de corte, da eficiência de conversão do alimento em produto: comparação entre grupos experimentais. **Revista Brasileira de Zootecnia**, Brasília, v. 41, n. 1, p. 138-146, jan. 2012.

DRAPER, N. R.; SMITH, N. **Applied regression analysis**. 3th ed. New York: John Wiley & Sons, 1998. 706 p.

DREHMER, A. M. F.; AMARANTE, C. V. T. Conservação pós-colheita de frutos de araçá-vermelho em função do estádio de maturação e temperatura de armazenamento. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 30, n. 2, p. 322-326, jun. 2008.

EMILIANO, P. C.; VIVANCO, M. J. F.; MENEZES, F. S. de. Information criteria: how do they behave in different models? **Computational Statistics and Data Analysis**, Amsterdam, v. 69, p. 141-153, 2014.

FARIAS, J. F. de et al. Maturação e determinação do ponto de colheita de frutos de envira-caju. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 33, n. 3, p. 730-736, set. 2011.

FERNANDES, T. J. et. al. Seleção de modelos não lineares para a descrição das curvas de crescimento do fruto do cafeeiro. **Coffee Science**, Lavras, v. 9, n. 2, p. 204-212, abr./jun. 2014.

FERREIRA NETO, M. et al. Crescimento e produção de coqueiro Anão verde fertigado com nitrogênio e potássio. **Revista Brasileira de Engenharia Agrícola e Ambiental**, Campina Grande, v. 15, n. 7, p. 657-664, jul. 2011.

FERREIRA, D. F. **Estatística multivariada**. 2. ed. Lavras: Editora da UFLA, 2011. 676 p.

FORNI, S. **Análise da curva de crescimento de bovinos da raça Nelore, utilizando modelos não-lineares em análises bayesianas**. 2007. 65 p. Tese (Doutorado em Zootecnia) - Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual de São Paulo, Jaboticabal, 2007.

GALLANT, A. R. **Nonlinear statistical models**. New York: John Wiley & Sons, 1987. 610 p.
GOMPERTZ, B. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. **Philosophical Transactions of the Royal Society of London**, London, n. 115, p. 513-583, Jan. 1825.

GRAYBILL, F.; IYER, H. **Regression analysis:** concepts and applications. Belmont: Duxbury Press, 1994. 648 p.

GURJÃO, K. C. de O. et al. Desenvolvimento de frutos e sementes de tamarindo. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 28, n. 3, p. 351-354, dez. 2006.

HOLANDA, J. S. de et al. **Tecnologias para produção intensiva de coco anão**. Natal: EMPARN, 2007. 40 p. (Boletim de Pesquisa, 34).

HUNT, R. **Plant growth analysis**. London: Editora Edward Arnold, 1978. 67 p. (Studies in Biology, 96).

JESUS, N. de; MARTINS, A. B. G.; BARBOSA, J. C. Desenvolvimento de frutos de Longan (Dimocarpus longan lour) na região de Jaboticabal-SP. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 30, n. 1, p. 159-164, mar. 2008.

LACERDA, M. A. D.; LACERDA, R. D.; ASSIS, P. C. de O. A participação da fruticultura no agronegócio brasileiro. **Revista de Biologia e Ciências da Terra**, Campina Grande, v. 4, n. 1, p. 01-09, jan./jun. 2004.

LOAIZA-ECHEVERRI, A. M. Utilização de modelos não-lineares para descrever o crescimento do perímetro escrotal em ruminantes. **Revista Veterinária e Zootecnia**, Botucatu, v. 5, n. 2, p. 9-23, jun. 2011.

LOIOLA, C. M. **Comportamento de cultivares de coqueiro (Cocos nuciferaL.) em diferentes condições agroecológicas dos tabuleiros costeiros do nordeste brasileiro**. 2009. 74 p. Dissertação (Mestrado em Agroecossistema) - Universidade Federal de Sergipe, São Cristóvão, 2009.

MACIEL, V. T. et al. Caracterização física dos frutos de seis cultivares de coqueiro anão verde em diferentes estádios de desenvolvimento. **Revista Brasileira de Ciências Agrárias**, Recife, v. 4, n. 4, p. 395-398, out./dez. 2009.

MACIEL, V. T. **Caracterização física, físico-química e enzimática de frutos de seis cultivares de coqueiro anão em diferentes estádios de desenvolvimento**. 2008. 103 p. Dissertação (Mestrado em Fitotecnia) - Universidade Federal do Ceará, Fortaleza, 2008.

MARTINS FILHO, S. et al. Abordagem Bayesiana das curvas de crescimento de duas cultivares de feijoeiro. **Ciência Rural**, Santa Maria, v. 38, n. 6, p. 1516-1521, set. 2008.

MAZZINI, A. R. de A. et al. Curva de crescimento de novilhos Hereford: heterocedasticidade e resíduos autorregressivos. **Ciência Rural**, Santa Maria, v. 35, n. 2, p. 422-427, mar./abr. 2005.

MOTULSKY, H.; CHRISTOPOULOS, A. **Fitting models to biological data using linear and nonlinear regression:** a practical guide to curve fitting. 4th ed. San Diego: GraphPad Software, 2003. 351 p.

PEIXOTO, C. P. et al. Análise de crescimento de diferentes genótipos de citros cultivados sob déficit hídrico. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 28, n. 3, p. 439-443, dez. 2006.

PEREIRA, M. C. T. et al. Crescimento e produção de primeiro ciclo da bananeira ‘Prata Anã’(AAB) em sete espaçamentos. **Pesquisa Agropecuária Brasileira**, Brasília, v. 35, n. 7, p. 1377-1387, jul. 2000.

PRADO, T. K. L. do et al. A. Ajuste do modelo logístico na descrição do crescimento de frutos de coqueiro anão por meio de algoritmos iterativos MCMC. **Revista Brasileira de Biometria**, São Paulo, v. 31, n. 2, p. 216-232, abr./jun. 2013.

PRADO, T. K. L. do; SAVIAN, T. V.; MUNIZ, J. A. Ajuste dos modelos Gompertz e Logístico aos dados de crescimento de frutos de coqueiro anão verde. **Ciência Rural**, Santa Maria, v. 43, n. 5, p. 803-809, maio 2013.

RESENDE, J. M. et al. Colheita e manuseio póscolheita. In: ARAGÃO, W. M. (Ed.). **Coco:** pós-colheita. Brasília: Embrapa Informação Tecnologia, 2002. p. 11-18. (Série Frutas do Brasil, 29).

RIBEIRO, T. D. **Modelos de Regressão não linear na descrição da cinética de secagem de polpa de jabuticaba**. 2015. 95 p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) – Universidade Federal de Lavras, Lavras, 2015.

SEARLE, S. R. **Linear models**. New York: John Wiley & Sons, 1971. 532 p.

SEBER, G. A. F.; WILD, C. J. **Nonlinear regression**. Hoboken: John Wiley & Sons, 2003. 792 p.

SILVA, D. L. V. da et al. Características físicas, físico-químicas e sensoriais da água de frutos de coqueiro anão verde oriundo de produção convencional e orgânica. **Ciência e Agrotecnologia**, Lavras, v. 33, n. 4, p. 1079-1084, jul./ago. 2009.

SILVA, F. de L. et al. Curvas de crescimento em vacas de corte de diferentes tipos biológicos. **Pesquisa Agropecuária Brasileira**, Brasília, v. 46, n. 3, p. 262-271, mar. 2011.

SILVA, L. R. da et al. Características físicas e físico-químicas da água de frutos de coqueiro anão verde. **Revista Brasileira de Tecnologia Agroindustrial**, Ponta Grossa, v. 7, n. 2, p. 1022-1032, 2013.

SOUZA, I. F. et al. Fitting nonlinear autoregressive models to describe coffee seed germination. **Ciência Rural**, Santa Maria, v. 44, n. 11, p. 2016-2021, nov. 2014.

SOUZA, L. de A. et al. Curvas de crescimento em bovinos da raça Indubrasil criados no Estado do Sergipe. **Revista Ciência Agronômica**, Fortaleza, v. 41, n. 4, p. 671-676, nov./dez. 2010.
STATISTICAL ANALYSIS SYSTEM INSTITUTE. **SAS procedures guide for computers**: volume 3. 6th ed. Cary: SAS Institute, 2014. 373 p.

STRATHE, A. B. et al. A multivariate nonlinear mixed effects method for analyzing energy partitioning in growing pigs. **Journal of Animal Science**, Champaign, v. 88, n. 7, p. 2361-2372, Mar. 2010.

SUBRAMANIAN, A.; SILVA, L. B. da; COUTINHO, A. S. Aplicação de método e técnica multivariados para previsão de variáveis termoambientais e perceptivas. **Production**, São Paulo, v. 17, n. 1, p. 52-70, abr. 2007.

TERRA, M. F.; MUNIZ, J. A.; SAVIAN, T. V. Ajuste dos modelos Logístico e Gompertz aos dados de crescimento de frutos da tamareira-anã (*Phoenix roebelenii* O'BRIEN). **Magistra**, Cruz das Almas, v. 22, n. 1, p. 1-7, 2010.

URCHEI, M. A.; RODRIGUES, J. D.; STONE, L. F. Análise de crescimento de duas cultivares de feijoeiro sob irrigação, em plantio direto e preparo convencional. **Pesquisa Agropecuária Brasileira**, Brasília, v. 35, n. 3, p. 497-506, mar. 2000.

XU, S. et al. Multivariate nonlinear mixed model to analyze longitudinal image data: MRI study of early brain development. In: CONFERENCE ON COMPUTER IEEE COMPUTER SOCIETY, 2008, Anchorage. **Proceedings...** Anchorage: IEEE, 2008. p. 1-8.

ANEXOS

ANEXO A – Rotina no SAS

Obtidos os ajustes dos modelos não lineares univariados para cada variável resposta, o vetor de resíduos em cada caso foi utilizado para obtenção das matrizes de covariâncias e o fator de Cholesky, como segue:

```
proc iml;
E1={0.3144     0.1773  -0.2201 -0.1620 -0.5000        -0.0132     -0.3480          -
0.4003          -0.0494  -0.1547 0.284            -0.0370 -0.0027          0.1857
0.2612  -0.4527 0.7224           1.3480 -0.1588   0.2674           0.1367           -
0.1668             -0.6164        -0.2279 -0.5958         -0.0059,
0.2192            0.1014           -0.2644            -0.2875          -0.4453          -
0.3276 -0.4317  -0.3658 -0.4661 -0.0665 0.0721          0.2162 0.1047
0.9262
1.0839          -0.1797  0.8333            0.7784            -0.4446          -0.2067 -
0.2063  0.1259            -0.3545 -0.1334 -0.6132      -0.3804};
a=E1;
print a;
run;
sigma=(E1*t(E1))/26; print sigma;
p=half(ginv(sigma));print p;
```

Os vetores de resíduos acima correspondem ao modelo gompertz univariado para as duas variáveis, DEL e DET. O ajuste do modelo multivariado depende da matriz de Cholesky e é dado por:

```
data work01; set teste1;
p1= 3.4905518; p2= -2.261837; Y=p1*DEL+p2*DET; output;
p1=0; p2= 2.1842764; Y=p1*DEL+p2*DET; output;delete;

proc model data=work01 maxiter=1000;
 Y=p1*(a1/(1+ b1 * exp(-k1*x)))+p2*(a2/(1+ b2 * exp(-k2*x)));
  parms  a1=20 b1=10 k1=0.05 a2=15 b2=5 k2=0.05;
  fit Y/  outall out=work02 dw=4 prl=wald dwprob;
%ar(Y,p);
run;
```

Para estrutura de erros independentes tirar a linha de comando %ar(Y,p), em que p é a ordem para a estrutura de autocorrelação.