



TALES JESUS FERNANDES

**MODELOS DUPLO SIGMOIDAIS PARA O
CRESCIMENTO DE FRUTOS DO CAFEIEIRO**

LAVRAS - MG

2016

TALES JESUS FERNANDES

**MODELOS DUPLO SIGMOIDAIS PARA O CRESCIMENTO DE
FRUTOS DO CAFEIRO**

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

Orientador
Dr. Joel Augusto Muniz

**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)**

Fernandes, Tales Jesus

Modelos duplo sigmoidais para o crescimento de frutos do
cafeeiro. / Tales Jesus Fernandes. – Lavras : UFLA, 2016.

89 p.

Tese (doutorado) – Universidade Federal de Lavras, 2016.

Orientador: Joel Augusto Muniz.

Bibliografia.

1. Bayesiana. 2. Café. 3. Duplo Logístico. 4. Reamostragem
por importância. 5. Regressão não linear. I. Universidade Federal de
Lavras. II. Título.

TALES JESUS FERNANDES

**MODELOS DUPLO SIGMOIDAIS PARA O CRESCIMENTO DE
FRUTOS DO CAFEIRO**

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

APROVADA em 29 de Abril de 2016.

Prof. Dr. Júlio Sílvio de Souza Bueno Filho	UFLA
Prof. Dra. Carla Regina Guimarães Brighenti	UFSJ
Prof. Dra. Izabela Regina Cardoso de Oliveira	UFLA
Prof. Dr. Augusto Ramalho de Morais	UFLA

Dr. Joel Augusto Muniz
Orientador

**LAVRAS - MG
2016**

À minha esposa Adriele pela paciência e companheirismo em todos os momentos.

Aos meus pais, Iris e Sebastião, pelas palavras de incentivo e confiança.

Ao meu irmão, Waynis, pelo apoio e admiração.

DEDICO.

AGRADECIMENTOS

Primeiramente, à DEUS, pois toda vitória só é possível com Ele.

Ao professor Joel Augusto Muniz, pela confiança, pela orientação e pelas experiências compartilhadas.

Ao professor Júlio Sílvio de Souza Bueno Filho pelas constantes sugestões e apoio no desenvolvimento deste trabalho.

Aos demais professores do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, agora meus colegas, pelo conhecimento transmitido e sua contribuição para a minha formação pessoal, acadêmica e profissional.

Ao professor Eric Batista Ferreira por ter me incentivado e ajudado com os primeiros passos na Estatística.

Aos professores membros da banca, de qualificação e defesa, pelas sugestões que contribuíram e muito para enriquecer este trabalho.

À Universidade Federal de Lavras, por me oferecer ensino gratuito de qualidade, e agora por fornecer a oportunidade de exercer minha vocação na transmissão de conhecimento.

Aos meus primos e meus amigos pelo companheirismo, pela admiração e pelos momentos de descontração, essenciais para chegar até aqui.

À minha esposa, Adrielle, por estar sempre ao meu lado em todos os momentos, com suas palavras de confiança e incentivo, sempre disposta a ajudar.

À minha família, principalmente aos meus pais, pois sempre me apoiaram e incentivaram, obrigado por me prepararem tão bem para a vida. Seu amor, carinho e orações foram essenciais para concluir mais esta etapa.

“Estude enquanto eles dormem,
trabalhe enquanto eles se divertem,
persista enquanto eles descansam.
E então, viva o que eles sonham.”
Autor desconhecido

“Mas vocês têm de saber que não
há um prêmio no final do caminho.
O prêmio é o caminho mesmo.”
José Alberto Mujica Cordano

“Ninguém se sente feliz vivendo o
tempo presente. Só se sabe que foi
feliz lembrando do tempo ausente.”
Clélia Coelho Pedroso

RESUMO GERAL

O café é um dos principais produtos da agricultura brasileira e o país é atualmente o maior produtor e exportador mundial. Conhecer o padrão de crescimento de um fruto pode auxiliar no desenvolvimento da cultura indicando por exemplo, as épocas de maior ganho de massa do fruto e seu ponto ótimo de colheita, fundamentais para melhorar o manejo e a qualidade do café. De acordo com alguns autores o fruto do café apresenta um padrão de crescimento duplo sigmoidal, no entanto, assim como em outros frutos que também apresentam tal padrão de crescimento, os autores não estimam parâmetros de modelos de regressão para descrever tal curva. Os objetivos deste trabalho foram: *i)* verificar se o padrão de crescimento de frutos do cafeeiro é duplo sigmoidal; *ii)* propor uma modificação na metodologia de reamostragem por importância ponderada de modo a facilitar, ainda mais, o seu uso na estimação de parâmetros e selecionar o modelo não linear duplo sigmoidal mais adequado para descrever o crescimento de frutos do cafeeiro; *iii)* estudar o efeito da distribuição espacial da lavoura na curva de crescimento de frutos do cafeeiro. No primeiro artigo o objetivo foi verificar se o padrão de crescimento de frutos do cafeeiro realmente possui aspecto duplo sigmoidal de crescimento. Os modelos duplo Gompertz e duplo Logístico apresentaram qualidade de ajuste significativamente superior aos modelos de sigmóides simples confirmando que o padrão de crescimento de frutos do cafeeiro é realmente duplo sigmoidal. No segundo artigo propõe-se considerar uma aproximação da verossimilhança como a distribuição candidata da metodologia de reamostragem por importância ponderada, objetivando facilitar o processo de obtenção de amostras das distribuições marginais de cada parâmetro. Tal técnica foi eficiente, pois utilizou baixo esforço computacional e forneceu parâmetros com

interpretação prática, logo pode ser utilizada para estimar parâmetros de curvas de crescimento duplo sigmoidais. Dentre os modelos estudados, selecionou-se o modelo duplo Logístico como o mais adequado para descrever a curva de crescimento de frutos do cafeeiro. No terceiro artigo objetivou-se verificar a influência dos diferentes alinhamentos de plantio e das faces de exposição ao sol na curva de crescimento dos frutos. Percebeu-se diferença entre as taxas de crescimento nas duas etapas de uma mesma curva ao longo do desenvolvimento do fruto, independente da face. Embora já tenha sido observado que existe diferença na qualidade e na produtividade entre faces e alinhamentos, neste trabalho, concluiu-se que estes fatores não influenciam a curva de crescimento dos frutos do cafeeiro.

Palavras-chave: Bayesiana; Café; Duplo Logístico; Reamostragem por importância; Regressão não linear.

GENERAL ABSTRACT

Coffee is one of the main products of Brazilian agriculture, the country is currently the largest producer and exporter. Knowing the growth pattern of a fruit can assist in the development of culture indicating for example, the times of increased fruit weight and its optimum harvest, essential to improve the management and quality of coffee. Some authors indicate that the growth curve of the coffee fruit has a double sigmoid shape. However, it consists of just a visual observation without exploring the use of regression models. The aims of this study were: *i)* determine if the growth pattern of the coffee fruit is really double sigmoidal; *ii)* to propose a new approach in weighted importance re-sampling to estimate the parameters of regression models and select the most suitable double sigmoidal model to describe the growth of coffee fruits; *iii)* to study the spatial distribution effect of the crop in the growth curve of coffee fruits. In the first article the aim was determine if the growth pattern of the coffee fruit is really double sigmoidal. The models double Gompertz and double Logistic showed significantly superior fit to models of simple sigmoid confirming that the standard of coffee fruits growth is really double sigmoidal. In the second article we propose to consider an approximation of the likelihood as the candidate distribution of the weighted importance resampling, aiming to facilitate the process of obtaining samples of marginal distributions of each parameter. This technique was effective since it provided parameters with practical interpretation and low computational effort, therefore, it can be used to estimate parameters of double sigmoidal growth curves. The nonlinear model double Logistic was the most appropriate to describe the growth curve of coffee fruits. In the third article aimed to verify the influence of different planting alignments and sun exposure faces in the fruits growth curve.

A difference between the growth rates in the two stages of fruit development was identified, regardless the side. Although it has been proven differences in productivity and quality of coffee, there was no difference between the growth curves in the different planting alignments herein studied.

Keywords: Bayesian; Coffee; Double Logistic; Importance resampling; Nonlinear regression.

SUMÁRIO

1	CAPÍTULO 1 Modelos duplo sigmoidais na descrição da curva de crescimento do fruto do cafeeiro.	12
1.1	INTRODUÇÃO	14
1.2	MATERIAL E MÉTODOS	19
1.3	RESULTADOS E DISCUSSÃO	23
1.4	CONCLUSÕES	29
	REFERÊNCIAS	30
2	CAPÍTULO 2 Estimação de parâmetros de modelos duplo sigmoidais: Uma abordagem bayesiana utilizando reamostragem por importância ponderada.	34
2.1	INTRODUÇÃO	36
2.2	MATERIAL E MÉTODOS	41
2.3	RESULTADOS E DISCUSSÃO	53
2.4	CONCLUSÕES	63
	REFERÊNCIAS	64
3	CAPÍTULO 3 Avaliação do efeito dos alinhamentos de plantio na curva de crescimento do fruto do cafeeiro.	68
3.1	INTRODUÇÃO	70
3.2	MATERIAL E MÉTODOS	72
3.3	RESULTADOS E DISCUSSÃO	76
3.4	CONCLUSÕES	85
	REFERÊNCIAS	86

1 CAPÍTULO 1 Modelos duplo sigmoidais na descrição da curva de crescimento do fruto do cafeeiro.

RESUMO

Uma das formas de auxiliar na melhoria da produção cafeeira e, principalmente, da qualidade da bebida é conhecer o padrão de crescimento dos frutos, possibilitando assim um melhor manejo da cultura. Alguns autores indicam que a curva de crescimento do fruto do cafeeiro tem um formato duplo sigmoidal, no entanto tal identificação é feita apenas visualmente sem explorar o uso de modelos de regressão. O objetivo deste trabalho foi verificar se o padrão de crescimento do fruto do cafeeiro, considerando acúmulo de massa fresca em função do tempo, é realmente duplo sigmoidal e selecionar o modelo não linear mais indicado para descrever tal comportamento. Os dados utilizados foram extraídos de Cunha e Volpe (2011) obtidos em um experimento com a cultivar Obatã IAC 1669-20. Foram comparados os ajustes fornecidos pelos modelos Logístico e Gompertz em suas versões simples e duplo. A estimação dos parâmetros foi feita pelo método dos mínimos quadrados utilizando o algoritmo de Gauss-Newton implementado na função *nls* do software R. Pode-se concluir que o padrão de crescimento do fruto do cafeeiro, em acúmulo de massa fresca, é duplo sigmoidal. Os modelos duplo Gompertz e duplo Logístico alcançaram ótimos resultados, ambos com coeficiente de determinação ajustado superior a 99%, com uma ligeira superioridade do modelo duplo Logístico.

Palavras-chave: Café; Duplo Gompertz; Duplo Logístico; Não linear.

ABSTRACT

Knowing the growth pattern of coffee fruits is a way to assist its production improvement and, especially, the drink quality in order to obtain better crop management. Some authors indicate that the growth curve of the coffee fruit has a double sigmoid shape. However, it consists of just a visual observation without exploring the use of regression models. The aim of this study was to determine if the growth pattern of the coffee fruit, considering the accumulation fresh weight as a function of time, is really double sigmoidal and select the most suitable non-linear model to describe this behavior. The data used in this work is from Cunha and Volpe (2011), who studied the Obatã IAC 1669-20 crops. Adjustments obtained from the Logistic and Gompertz models in their single and dual versions were compared. The parameter estimation was made by the least squares method using the Gauss-Newton algorithm implemented in the `nls` function of the R Statistical software. It can be concluded that the growth pattern of the coffee fruit in fresh weight accumulation is double sigmoidal. The double Gompertz and double Logistic models achieved excellent results, both with higher coefficient of determination that was superiorly adjusted to 99%, with a slight superiority of the double Logistic model.

Keywords: Coffee; Double Gompertz; Double Logistic; Nonlinear.

1.1 INTRODUÇÃO

Historicamente o café é um dos principais produtos da agricultura brasileira, com relevante importância econômica e social. O país é atualmente o maior produtor e exportador mundial de café e segundo maior mercado consumidor do produto. Pertencente ao gênero *Coffea*, o café possui diversas espécies no mundo, sendo que apenas duas apresentam importância econômica, o *Coffea arabica* e *Coffea canephora*, popularmente conhecidos como café arábica e conilon, respectivamente. De acordo com a Companhia Nacional de Abastecimento - CONAB (2015), devido à melhor qualidade da bebida e, conseqüentemente, ao maior valor de mercado, o café arábica corresponde a mais de 70% da produção nacional e mundial.

A preocupação em produzir café de qualidade e também em aumentar a produtividade tem levado os produtores a investirem cada vez mais em processos que envolvem a melhoria do manejo da cultura (SILVA et al., 2011). Nesse sentido, estudos que visam descrever as curvas de crescimento dos frutos do cafeeiro podem auxiliar o produtor na identificação de possíveis pontos de intervenção durante o desenvolvimento do fruto, tais como a época de maior ganho de massa e o momento mais adequado para a adubação, obtendo assim uma melhor qualidade final do produto.

Um problema particularmente intrigante e útil para pesquisadores e produtores é prever o tamanho e, ou, peso médio dos frutos na época da colheita, com base em informações disponíveis em um momento anterior do ano agrícola, possivelmente no início do ciclo produtivo. Principalmente no caso do café que é comercializado em sacas de 60kg, daí a importância de se conhecer o peso final do produto. Naturalmente, tal informação depende de uma série de fatores ambientais e de manejo, que podem variar de ano para ano, mas geralmente

são bem estabelecidos no início do ciclo produtivo. Vários métodos podem ser utilizados neste contexto, sendo que um dos principais é a determinação da curva de crescimento do fruto (BEBBINGTON et al., 2009).

Segundo Coombe (1976), estudar o padrão de crescimento de um fruto auxilia no estabelecimento de índices de maturidade, possibilitando assim a determinação do ponto ótimo de colheita. Este ponto, que é um dos principais fatores de perda pós-colheita, é determinado, na maioria das vezes, de maneira subjetiva de acordo com Cavalini et al. (2006). Além disso, o conhecimento das curvas de crescimento do fruto do cafeeiro auxilia na identificação das diferentes fases fenológicas envolvidas em seu desenvolvimento.

O cafeeiro leva dois anos para completar o ciclo fenológico, o qual possui uma sucessão de fases vegetativas e reprodutivas, diferentemente da maioria das plantas que emitem as inflorescências na primavera e frutificam no mesmo ano. De acordo com Ságio (2009), o primeiro ano fenológico inicia com o crescimento vegetativo e formação das gemas foliares, posteriormente surgem as gemas florais passando pela indução, maturação e dormência. O segundo ano fenológico inicia-se com a florada que acontece no reinício das chuvas da primavera seguida pelo aparecimento do fruto que ocorre entre 6 a 8 semanas após a queda das flores (CAMARGO; CAMARGO, 2001; RENA; MAESTRI, 1985). A partir deste ponto, autores como Castro e Marraccini (2006), Geromel et al. (2006), Nunes et al. (2010) e Ságio (2009) classificam o desenvolvimento do fruto em cinco estágios da seguinte maneira:

- I- **chumbinho** - é um estágio caracterizado pela baixa taxa de crescimento, que ocorre nos dois primeiros meses após a antese, mais precisamente até a 10^a semana, estando os frutos bem pequenos ocorre uma intensa divisão celular nesse período;

- II- **crescimento rápido** - este estágio é caracterizado pela rápida expansão das células dos frutos, quando ocorre um aumento abrupto de tamanho e de massa fresca, geralmente, entre os meses de dezembro e janeiro;
- III- **crescimento lento** - neste estágio acontece uma desaceleração do crescimento, que pode se estender por até duas semanas, na qual o endosperma transforma-se a partir de um líquido para o estado sólido promovendo a formação das sementes do café e o volume final do fruto é atingido;
- IV- **granação** - ocorre geralmente entre a 17^a e a 28^a semana após a antese, em pleno verão, quando a matéria seca e fresca do fruto voltam a aumentar exclusivamente por causa do aumento da massa do pericarpo;
- V- **maturação** - neste, que é o último estágio, ocorrem mudanças na coloração externa do fruto, se tornando vermelho ou amarelo. A maturação pode se estender por um período de aproximadamente 10 semanas, da 24^a a 34^a semana (geralmente, nos meses de abril a junho). Ao fim deste estágio, ocorre perda de água e o fruto está completamente formado, suas sementes estão prontas e, portanto, apto a ser colhido.

O tempo de duração de cada estágio pode sofrer variações em função das espécies e, principalmente, das condições climáticas da região, sendo que cada estágio tem funções fisiológicas e metabólicas próprias, essenciais à formação final do fruto (LAVIOLA et al., 2007). O fruto do café tem formato ovóide e, geralmente, contém duas sementes, mas podem também ocorrer frutos com apenas uma semente ocupando todo o espaço do ovário, denominados grão moça. A Figura 1.1 apresenta um esquema do desenvolvimento do fruto, identificando os cinco estágios descritos anteriormente.

Ao longo destes estágios, acompanhando a massa fresca acumulada do

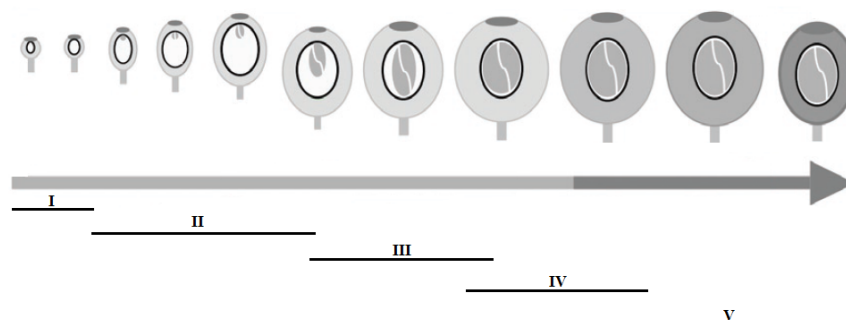


Figura 1.1 Estágios de desenvolvimento do fruto do cafeeiro, adaptada de Castro e Marraccini (2006).

fruto, identifica-se um padrão de crescimento duplo sigmoidal (BELAN et al., 2011; CASTRO; MARRACCINI, 2006; GEROMEL et al., 2006; LAVIOLA et al., 2007; MORAIS et al., 2008; RENA; MAESTRI, 1985). Mas devido a dificuldades envolvidas na estimação de parâmetros de curvas duplo sigmoidais, os autores apenas caracterizaram este padrão, sem efetivamente modelar tal crescimento e obter estimativas para os parâmetros de um modelo duplo sigmoidal.

Por serem capazes de descrever as mais diversas situações, apresentar melhor qualidade de ajuste mesmo utilizando poucos parâmetros e principalmente, apresentar interpretação prática para estes, os modelos de regressão não lineares são muito utilizados nas mais diversas áreas do conhecimento, principalmente no estudo de curvas de crescimento (CARNEIRO et al., 2014; FERNANDES et al., 2014; PEREIRA et al., 2014). Os modelos não lineares da família Richards são os mais indicados na literatura para a descrição de curvas de crescimento, são eles: Logístico, Gompertz, Brody e von Bertalanffy.

Ao se assumir que as curvas duplo sigmóide resultam de dois processos de crescimento consecutivos, então estas podem ser obtidas pela soma de duas funções de crescimento simples (ASCHONITIS et al., 2014; HAU; AMORIM;

BERGAMIM FILHO, 1993; MISCHAN et al., 2015). Desta forma, para descrever o crescimento duplo sigmoidal considera-se, em um único modelo, dois sigmóides simples, cada um com o objetivo de explicar uma determinada fase do crescimento. A escolha de qual modelo deve ser colocado em cada fase depende, obviamente, do padrão de crescimento em cada uma delas (SHEEHY; MITCHELL; FERRER, 2004).

Por serem os mais conhecidos e apresentarem boa capacidade de ajuste, os modelos duplo Logístico (Logístico + Logístico) e duplo Gompertz (Gompertz + Gompertz), aparecem com maior frequência neste tipo de estudo. Meyer (1994) propôs a curva duplo Logística para modelar o crescimento quando há duas fases distintas, cada uma com simetria em relação ao seu ponto de inflexão. Em presença de assimetria dentro de ambas fases, Hurwitz et al. (1991) foram um dos pioneiros a utilizar o modelo duplo Gompertz.

Grande parte dos trabalhos utilizando modelos duplo sigmóides está relacionada ao crescimento animal (HURWITZ et al., 1991; MENDES et al., 2008; MEYER, 1994; SANTORO et al., 2005; VASQUEZ et al., 2012). No entanto, em frutos de caroço, o crescimento deste notavelmente se sobrepõe ao crescimento da polpa em um primeiro instante, fazendo com que suas medidas em função do tempo apresentem dois pontos de inflexão caracterizando uma curva com formato de S duplo que pode ser descrita por dois sigmóides (COOMBE, 1976). Mas, a maioria dos autores que estudam o crescimento de frutos, apenas identifica visualmente tal padrão de crescimento sem proceder com o ajuste de modelos de regressão (BEBBINGTON et al., 2009; FAMIANI et al., 2012; NAVA et al., 2014; SANTOS et al., 2015; SILVA et al., 2013).

O objetivo deste trabalho foi verificar se o padrão de crescimento do fruto do cafeeiro, explicando a massa fresca acumulada em função do tempo, é duplo

sigmoidal e selecionar o modelo não linear mais indicado na descrição de tal comportamento. Foram comparados os modelos Logístico e Gompertz em suas versões simples e duplo.

1.2 MATERIAL E MÉTODOS

Os dados foram extraídos de Cunha e Volpe (2011) e correspondem aos resultados parciais de um experimento conduzido na Faculdade de Ciências Agrárias e Veterinárias, UNESP, Campus Jaboticabal, SP (altitude: 615m), com a cultivar Obatã IAC 1669-20, que pertence a espécie *C. arabica*. A lavoura estava implantada no alinhamento de plantio $51^\circ - 231^\circ$, o experimento foi realizado na safra 2007/2008, na qual as plantas do cafeeiro se encontravam com 5 anos de idade, altura média de 2m e espaçamento de 3,5x0,5m.

A lavoura experimental foi dividida em 4 parcelas com 100 plantas cada. As coletas foram feitas aleatoriamente nas parcelas, sendo colhidos 50 frutos da face norte e 50 da face sul. Acompanhou-se o desenvolvimento do fruto realizando 14 coletas com um espaçamento médio entre elas de 15 dias, sendo que a primeira foi obtida no dia 06/11/07, aos 96 dias após o florescimento (DAF) e a última no dia 21/05/08, aos 293 DAF. Os frutos eram acondicionados em sacos plásticos para serem pesados em balança analítica de 0,0001g de precisão e posteriormente obter a massa fresca média (em gramas).

Os modelos de regressão não linear utilizados foram o Logístico e Gompertz, em suas versões simples e duplo sigmoidal, pois, mesmo que as informações da literatura indiquem um padrão de crescimento duplo sigmoidal, é necessário verificar se ao dobrar o número de parâmetros do modelo obtém-se realmente um ganho na qualidade de ajuste, prezando assim pelo princípio da parcimônia. Existem diferentes parametrizações destes modelos, segundo Fernandes et al.

(2015) as que estão apresentadas na Tabela 1.1 devem ser preferidas por apresentar melhor comportamento em termos de não-linearidade.

Tabela 1.1 Equações dos modelos não lineares ajustados ao acúmulo de massa fresca dos frutos do cafeeiro.

Modelo	Expressão
Logístico	$y_i = \frac{a_2}{1+e^{k(b-x_i)}} + \varepsilon_i$
Duplo Logístico	$y_i = \frac{a_1}{1+e^{k_1(b_1-x_i)}} + \frac{a_2-a_1}{1+e^{k_2(b_2-x_i)}} + \varepsilon_i$
Gompertz	$y_i = a_2 e^{-e^{k(b-x_i)}} + \varepsilon_i$
Duplo Gompertz	$y_i = a_1 e^{-e^{k_1(b_1-x_i)}} + (a_2 - a_1) e^{-e^{k_2(b_2-x_i)}} + \varepsilon_i$

Na Tabela 1.1 tem-se que: i representa as observações longitudinais $i = 1, 2, \dots, 14$; y_i é a i -ésima observação da massa fresca média do fruto; x_i são as idades em dias após o florescimento (daf); a_2 é o valor final esperado para a massa fresca, ou peso a maturidade; a_1 é a massa fresca máxima esperada para a primeira fase do crescimento; k , k_1 e k_2 correspondem ao índice de maturidade das curvas e quanto maior estes índices menos tempo o fruto leva para atingir seu peso a maturidade; b , b_1 e b_2 são as abcissas dos pontos de inflexão das respectivas curvas, nesse ponto o fruto atinge o máximo de seu crescimento e começa a desacelerar até estabilizar; ε_i é o erro aleatório associado ao modelo, o qual pressupõe-se que seja independente e identicamente distribuído de modo que $\varepsilon \sim N(0, \sigma^2)$.

A estimação dos parâmetros dos modelos apresentados na Tabela 1.1 foi feita utilizando o método de mínimos quadrados, o qual consiste em minimizar a soma dos quadrados dos erros gerando o sistema de equações normais (SEN). Devido ao SEN ser não linear, este não possui forma fechada para as soluções e elas são obtidas por algoritmos iterativos, sendo que os mais conhecidos são: Gauss-Newton, Marquardt e Gradiente. Foi utilizado neste trabalho o método

iterativo de Gauss-Newton que está implementado na função *nls* da biblioteca *nlme* (PINHEIRO et al., 2015) do software R.

A qualidade do ajuste obtido pelos modelos foi avaliada com base nos seguintes critérios:

Coefficiente de determinação ajustado (R_{aj}^2):

Como os modelos estudados neste trabalho não apresentam o mesmo número de parâmetros, então deve ser utilizado o coeficiente de determinação ajustado que faz uma ponderação pelo número de parâmetros evitando assim que modelos com mais parâmetros apresentem maior valor de R_{aj}^2 sem realmente explicar mais sobre o fenômeno em questão. Sua expressão é dada por:

$$R_{aj}^2 = 1 - \frac{(1 - R^2)(n - 1)}{n - p} \quad (1.1)$$

em que R^2 é o coeficiente de determinação; n é número de observações e p o número de parâmetros do modelo. Maiores valores do coeficiente de determinação ajustado indicam melhores ajustes.

Crítério de informação de akaike (AIC):

Este critério é baseado na teoria da decisão, admite a existência de um modelo real que descreve os dados e é desconhecido, e tenta selecionar o modelo que mais se aproxima deste “modelo real”. O AIC é calculado pela seguinte expressão:

$$AIC = -2 \times \ln(L(\theta)) + 2p,$$

em que $\ln(L(\theta))$ é o valor do logaritmo de base natural do ponto de máximo da função de verossimilhança. Quanto menor o valor de AIC, melhor o ajuste

fornecido pelo modelo.

Desvio padrão residual (DPR):

Tal avaliador pode ser obtido pela seguinte expressão:

$$DPR = \sqrt{\frac{QME}{(n - p)}},$$

em que QME é o quadrado médio residual. Este valor é portanto proporcional ao quadrado médio do erro e, quanto menor, melhor é o ajuste fornecido pelo modelo.

Teste da Razão de Verossimilhança (TRV):

E finalmente para verificar a eficiência do modelo duplo sigmoidal em relação ao sigmóide simples utilizou-se o teste da razão de verossimilhança TRV. Este teste é indicado para comparar dois modelos hierarquicamente aninhados e as hipóteses em questão são:

H_0 : O modelo mais simples é o correto;

H_1 : O modelo com mais parâmetros é o correto.

O teste é calculado com base na seguinte expressão:

$$TRV = -2 * \ln \left(\frac{M_1}{M_2} \right) = 2[\ln(M_2) - \ln(M_1)] \sim \chi^2_{\nu}, \quad (1.2)$$

em que M_1 e M_2 são o ponto de máximo da função de verossimilhança do modelo mais simples e do mais completo, respectivamente. O teste é distribuído assintoticamente segundo uma distribuição de qui-quadrado com ν graus de liberdade, em que ν é a diferença do número de parâmetros entre o modelo mais completo e o mais simples. No entanto, de acordo com Regazzi e Silva (2004) para amostras

pequenas ($n < 120$) deve ser utilizada a estatística F ao invés da qui-quadrado, uma vez que as taxas de erro tipo I são menores considerando esta aproximação. Assim o critério de decisão do teste será com base no valor-p do teste F com o nível de significância de 5%. Portanto, rejeita-se H_0 caso $\text{valor} - p < 0,05$ sendo considerado o teste significativo, isto é, o modelo com mais parâmetros é o correto.

Toda a parte computacional deste trabalho, a estimação dos parâmetros, a obtenção dos critérios de qualidade de ajuste e construção de gráficos foi feita utilizando o software estatístico R (R CORE TEAM, 2015).

1.3 RESULTADOS E DISCUSSÃO

As estimativas dos critérios de qualidade de ajuste para os modelos Logístico simples e duplo apresentadas na Tabela 1.2 são todas favoráveis ao modelo duplo Logístico.

Tabela 1.2 Estimativas do coeficiente de determinação ajustado (R_{aj}^2), critério de informação de Akaike (AIC), desvio padrão residual (DPR) e valor-p do teste da razão de verossimilhanças (TRV) para as duas versões do modelo Logístico na descrição do crescimento do fruto do cafeeiro em massa fresca no alinhamento de plantio $51^\circ - 231^\circ$ em ambas as faces.

	R_{aj}^2	AIC	DPR	Valor-p TRV
FACE NORTE				
Logístico Simples	0,9461	-24,906	0,0842	0,0009
Duplo Logístico	0,9913	-46,14	0,0373	
FACE SUL				
Logístico Simples	0,9522	-27,383	0,0771	0,0001
Duplo Logístico	0,9954	-56,023	0,0262	

O teste de razão de verossimilhança foi significativo ($\text{valor} - p < 0,05$) indicando que o modelo com seis parâmetros é estatisticamente o mais adequado para descrever o crescimento em massa fresca do fruto do cafeeiro em função

do tempo. O modelo duplo Logístico apresentou os melhores resultados perante todos os demais critérios de qualidade inclusive com coeficiente de determinação ajustado (R_{aj}^2) superior a 99% em ambas as faces estudadas. Os resultados apresentados na Tabela 1.2 reforçam a ideia de que o crescimento do fruto do cafeeiro tem padrão de crescimento duplo sigmoidal assim como comentado por autores como Belan et al. (2011), Castro e Marraccini (2006) e Morais et al. (2008).

Nesse sentido, são apresentadas na Tabela 1.3 as estimativas, com os respectivos erros padrão, para os parâmetros do modelo duplo sigmoidal Logístico.

Tabela 1.3 Estimativas e erro padrão para os parâmetros do modelo duplo Logístico na descrição do crescimento do fruto do cafeeiro em massa fresca no alinhamento de plantio 51° - 231° em ambas as faces.

	a_1	b_1	k_1	a_2	b_2	k_2
FACE NORTE						
Estimativa	0,5972	146,1855	0,1492	1,0672	224,4171	0,0504
(erro padrão)	(0,0891)	(2,5901)	(0,0451)	(0,0473)	(8,8541)	(0,0212)
FACE SUL						
Estimativa	0,6554	150,8042	0,0982	1,0495	238,1695	0,0675
(erro padrão)	(0,0397)	(2,0182)	(0,0145)	(0,0301)	(4,6513)	(0,0211)

As estimativas dos parâmetros b_1 e b_2 indicam as idades (em dias) nas quais o fruto atingiu a taxa máxima de crescimento nas respectivas fases. Percebe-se que as estimativas para este parâmetro na face norte é menor que para a face sul sugerindo um padrão de crescimento diferente entre as faces da mesma planta. Tal fato também foi observado por Bicalho et al. (2005) e Custodio et al. (2013) que verificaram grandes diferenças na produtividade entre as faces da planta e afirmaram ainda, que a face com maior exposição ao sol apresenta maturação diferente com um acelerado processo de secagem dos frutos.

Como comentado por Bebbington et al. (2009), o peso à maturidade do

fruto é uma informação de extrema importância para os produtores, principalmente no caso do café por ser comercializado em sacas de 60kg. O parâmetro a_2 fornece esta informação e suas estimativas foram 1,06g e 1,04g, ao passo que a massa fresca máxima observada por Cunha e Volpe (2011) foi de 1,09g e 1,05g para as faces norte e sul respectivamente. Percebe-se aí uma tendência em subestimar a massa fresca máxima observada, mas tal fato é natural uma vez que após atingir o seu peso fresco máximo o fruto desidrata-se na própria planta e perde massa, passando a maturação seca (CUNHA; VOLPE, 2011). Assim, naturalmente o peso fresco à maturidade (a_2) deve mesmo ser menor que o máximo observado.

Foram estimados também os parâmetros para o modelo Gompertz cuja diferença em relação ao modelo Logístico é ser assimétrico ao ponto de inflexão. Os critérios de qualidade de ajuste e estimativas para os parâmetros deste modelo estão apresentados nas Tabelas 1.4 e 1.5 respectivamente.

Tabela 1.4 Estimativas do coeficiente de determinação ajustado (R_{aj}^2), critério de informação de Akaike (AIC), desvio padrão residual (DPR) e valor-p do teste da razão de verossimilhanças (TRV) para as duas versões do modelo Gompertz na descrição do crescimento do fruto do cafeeiro em massa fresca no alinhamento de plantio $51^\circ - 231^\circ$ em ambas as faces.

	R_{aj}^2	AIC	DPR	Valor-p TRV
FACE NORTE				
Gompertz Simples	0,9684	-30,857	0,0681	0,0085
Duplo Gompertz	0,9904	-44,279	0,0399	
FACE SUL				
Gompertz Simples	0,9737	-34,269	0,0603	0,0016
Duplo Gompertz	0,9946	-53,496	0,0287	

Novamente conforme pode ser observado na Tabela 1.4, o TRV foi significativo ($valor - p < 0,05$), indicando que o modelo com mais parâmetros é o mais adequado para descrever o crescimento do fruto do cafeeiro. Nos demais critérios,

todos apresentaram uma melhora significativa quando comparados os ajustes dos modelos duplo sigmóides em relação ao sigmóide simples. O coeficiente de determinação ajustado ($R_{a_j}^2$) fica acima dos 99%, ocorre uma diminuição no AIC e, principalmente, no DPR que fica reduzido praticamente a metade.

Tabela 1.5 Estimativas para os parâmetros do modelo duplo Gompertz na descrição do crescimento do fruto do cafeeiro em massa fresca no alinhamento de plantio 51° - 231° em ambas as faces.

	a_1	b_1	k_1	a_2	b_2	k_2
FACE NORTE						
Estimativa	0,6573	142,8341	0,0956	1,0774	220,1728	0,0392
(erro padrão)	(0,0642)	(2,0311)	(0,0333)	(0,0587)	(8,6195)	(0,0178)
FACE SUL						
Estimativa	0,7043	145,3653	0,0606	1,0643	234,3427	0,0473
(erro padrão)	(0,0393)	(1,8313)	(0,0103)	(0,0452)	(5,5911)	(0,0179)

Assim como no modelo Logístico as estimativas para os parâmetros b_1 e b_2 são menores para a face norte. Tal fato sugere que esta face apresenta uma curva de crescimento mais precoce, e sendo esta a face que recebe sol pela maior parte do tempo, estes resultados corroboram com os de Bicalho et al. (2005) e Custódio et al. (2013). Também pelo modelo Gompertz a estimativa do peso final a maturidade (a_2) está de acordo com os valores observados por Cunha e Volpe (2011).

Outra informação interessante é que tanto no modelo duplo Logístico quanto no modelo duplo Gompertz as estimativas para o parâmetro k_1 são maiores que para o k_2 . Indicando que o desenvolvimento do fruto do café é mais acelerado na primeira fase de crescimento (ou primeiro sigmóide), ao contrário, por exemplo, do que acontece com o pêssego que possui um crescimento mais notável na segunda fase como foi observado por Silva et al. (2013). Esta rápida expansão na primeira fase de desenvolvimento do fruto do café ocorre em coincidência com

o estágio de crescimento rápido da escala fenológica proposta por autores como Castro e Marraccini (2006) e Nunes et al. (2010).

Laviola et al. (2007) observaram que a taxa máxima de acúmulo de massa seca ocorreu aos 154 DAF na cultivar Catuaí IAC-44, sendo naturalmente, um pouco mais tardio que a taxa máxima de acúmulo de massa fresca estimadas neste trabalho, que foram aos 146 DAF e 150 DAF pelo modelo duplo Logístico e 142 DAF e 145 DAF pelo modelo duplo Gompertz, para as faces norte e sul respectivamente.

Como os modelos duplo sigmoidais apresentam os melhores ajustes, a Tabela 1.6 apresenta um resumo dos critérios de qualidade de ajuste para o modelo duplo Logístico e para o duplo Gompertz.

Tabela 1.6 Estimativas dos critérios de Qualidade para os modelos duplo Logístico e duplo Gompertz na descrição do crescimento do fruto do cafeeiro em massa fresca no alinhamento de plantio 51° - 231° em ambas as faces.

	R_{aj}^2	AIC	DPR
FACE NORTE			
Duplo Logístico	0,9913	-46,14	0,0373
Duplo Gompertz	0,9904	-44,279	0,0399
FACE SUL			
Duplo Logístico	0,9954	-56,023	0,0262
Duplo Gompertz	0,9946	-53,496	0,0287

Observando os resultados apresentados na Tabela 1.6, percebe-se que a qualidade de ajuste fornecida pelos dois modelos é semelhante, com os valores dos critérios muito próximos entre si. No entanto, deve ser destacado que mesmo estando próximos, o duplo Logístico apresenta melhores resultados perante todos os critérios de qualidade utilizados, com maior valor de R_{aj}^2 e menores valores de AIC e DPR. Desta forma, indica-se uma ligeira preferência pelo modelo

duplo Logístico na descrição do crescimento do fruto do cafeeiro, muito embora o modelo duplo Gompertz também apresente ótimos resultados como pode ser observado na Figura 1.2.

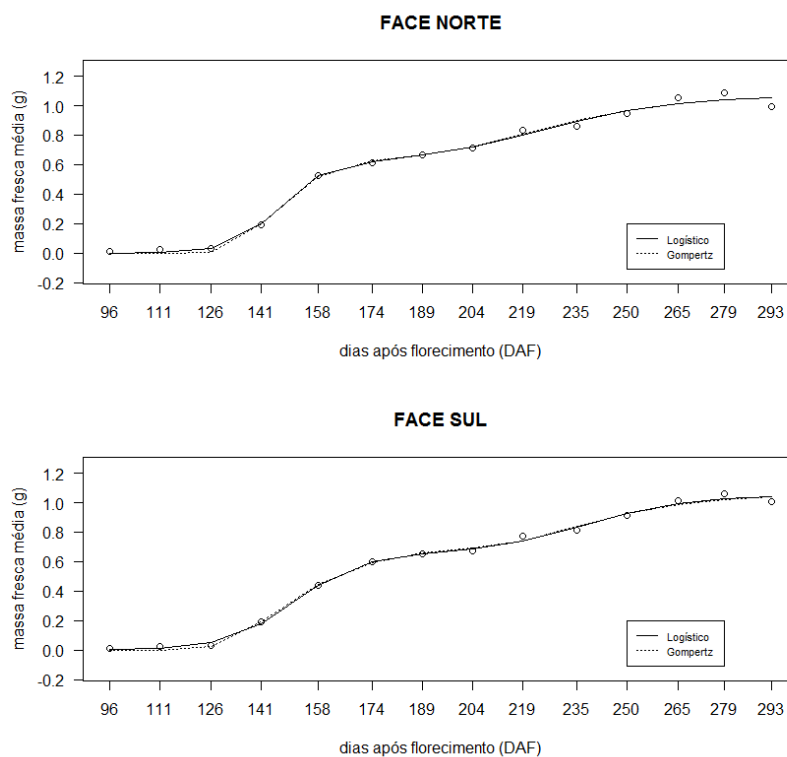


Figura 1.2 Ajuste dos modelos duplo sigmoidais Logístico e Gompertz aos dados de crescimento de frutos do cafeeiro das faces norte e sul do alinhamento de plantio 51° - 231°.

Percebe-se uma pequena diferença entre os modelos apenas no início do crescimento, até os 141 DAF, a partir daí os modelos duplo Logístico e duplo Gompertz são praticamente iguais.

1.4 CONCLUSÕES

O padrão de crescimento dos frutos do cafeeiro, considerando a massa fresca acumulada ao longo do tempo, é duplo sigmoidal.

Os modelos duplo Logístico e duplo Gompertz obtiveram ótimos desempenhos na descrição do padrão de crescimento do fruto do cafeeiro, com coeficiente de determinação ajustado superior a 99%. Todos os critérios de qualidade utilizados indicam uma ligeira superioridade do modelo duplo Logístico, de modo que este deve ser preferido em estudos sobre a modelagem da curva de crescimento de frutos do cafeeiro.

REFERÊNCIAS

- ASCHONITIS, V. G. et al. A review and synthesis of bivariate non-linear models to describe the relative variation of ecological, biological and environmental parameters. **Environmental Modelig and Assessment**, London, v. 20, p. 169-182, July 2014.
- BEBBINGTON, M. et al. Dynamics and phases of kiwifruit (*Actinidia deliciosa*) growth curves. **New Zealand Journal of Crop and Horticultural Science**, Wellington, v. 37, n. 3, p. 179-188, July 2009.
- BELAN, L. L. et al. Aspectos fisiológicos do cafeeiro conilon: uma abordagem sistemática. **Nucleus**, Ituverava, v. 8, n. 1, p. 225-240, abr. 2011.
- BICALHO, G. O. D. et al. Direcionamento das linhas de plantio em diferentes orientações cardeais e seus reflexos sobre a produtividade de cafeeiros. In: SIMPÓSIO DE PESQUISA DOS CAFÉS DO BRASIL, 1., 2005, Londrina. **Anais...** Brasília: EMBRAPA Café, 2005. Disponível em: <<http://www.sbicafe.ufv.br/handle/10820/1446>>. Acesso em: 5 jan. 2016.
- CAMARGO, A. P.; CAMARGO, M. B. P. Definição e esquematização das fases fenológicas do cafeeiro arábica nas condições tropicais do Brasil. **Bragantia**, Campinas, v. 60, n. 1, p. 65-68, mar. 2001.
- CARNEIRO, A. P. S. et al. Identidade de modelos não lineares para comparar curvas de crescimento de bovinos da raça Tabapuã. **Pesquisa Agropecuária Brasileira**, Brasília, v. 49, n. 1, p. 57-62, jan. 2014.
- CASTRO, R. D.; MARRACCINI, P. Cytology, biochemistry and molecular changes during coffee fruit development. **Brazilian Journal of Plant Physiology**, Campos dos Goytacazes, v. 18, n. 1, p. 175-199, Jan./Mar. 2006.
- CAVALINI, F. C. et al. Maturity indexes for “Kumagai” and “Paluma” guavas. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 28, n. 2, p. 176-179, ago. 2006.
- COMPANHIA NACIONAL DE ABASTECIMENTO. **Acompanhamento da safra brasileira: café**, v. 2, safra 2015, n. 2, segundo levantamento. Brasília, 2015. 59 p.

- COOMBE, B. G. The development of fleshy fruits. **Annual Review of Plant Physiology**, Palo Alto, v. 27, p. 507-528, 1976.
- CUNHA, A. R.; VOLPE, C. A. Curvas de crescimento do fruto de cafeeiro cv. Obatã IAC 1669-20 em diferentes alinhamentos de plantio. **Semina: Ciências Agrárias**, Londrina, v. 32, n. 1, p. 49-62, jan./mar. 2011.
- CUSTODIO, A. A. P. et al. Características vegetativas e produtivas de cafeeiros em diferentes supressões de irrigação e faces de exposição. **Coffee Science**, Lavras, v. 8, n. 4, p. 411-422, out./dez. 2013.
- FAMIANI, F. et al. Development and metabolism of the fruit and seed of the Japanese plum Ozark premier (*Rosaceae*). **Journal of Plant Physiology**, Stuttgart, v. 169, p. 551-560, Apr. 2012.
- FERNANDES, T. J. et al. Parameterization effects in nonlinear models to describe growth curves. **Acta Scientiarum. Technology**, Maringá, v. 37, n. 4, p. 397-402, Oct./Dec. 2015.
- FERNANDES, T. J. et al. Seleção de modelos não lineares para descrição das curvas de crescimento do fruto do cafeeiro. **Coffee Science**, Lavras, v. 9, n. 2, p. 207-215, abr./jun. 2014.
- GEROMEL, C. et al. Biochemical and genomic analysis of sucrose metabolism during coffee (*Coffea arabica*) fruit development. **Journal of Experimental Botany**, Lancaster, v. 57, n. 12, p. 3243-3258, Aug. 2006.
- HAU, B.; AMORIM, L.; BERGAMIN FILHO, A. Mathematical functions to describe disease progress curves of double sigmoid pattern. **Phytopathology**, Saint Paul, v. 83, n. 9, p. 928-932, 1993.
- HURWITZ, S. et al. Characterization of growth and development of Male British United Turkeys. **Poultry Science**, Champaign, v. 70, p. 2419-2424, 1991.
- LAVIOLA, B. G. et al. Acúmulo de nutrientes em frutos de cafeeiro em quatro altitudes de cultivo: cálcio, magnésio e enxofre. **Revista Brasileira de Ciência do Solo**, Campinas, v. 31, p. 1451-1462, maio 2007.
- MENDES, P. N. et al. Modelo logístico difásico no estudo do crescimento de fêmeas da raça Hereford. **Ciência Rural**, Santa Maria, v. 38, n. 7, p. 1984-1990, out. 2008.

MEYER, P. Bi-logistic growth. **Technological Forecasting and Social Change**, New York, v. 47, p. 89-102, 1994.

MISCHAN, M. M. et al. Inflection and stability points of diphasic logistic analysis of growth. **Scientia Agricola**, Piracicaba, v. 72, n. 3, p. 215-220, May/June 2015.

MORAIS, H. et al. Escala fenológica detalhada da fase reprodutiva de *Coffea arabica*. **Bragantia**, Campinas, v. 67, n. 1, p. 693-699, abr. 2008.

NAVA, A. D. et al. Growth kinetics of vegetative and reproductive organs of guava (*Psidium guajava* L.) in Iguala Guerrero, Mexico. **Agricultural Sciences**, Cambridge, v. 5, p. 1468-1475, Dec. 2014.

NUNES, F. L. et al. Modelos agrometeorológicos de estimativa da duração do estágio floração-maturação para três cultivares de café arábica. **Bragantia**, Campinas, v. 69, n. 4, p. 1011-1018, dez. 2010.

PEREIRA, A. A. et al. Descrição do crescimento vegetativo do cafeeiro cultivar Rubi MG 1192, utilizando modelos de regressão. **Coffee Science**, Lavras, v. 9, n. 2, p. 236-271, abr./jun. 2014.

PINHEIRO J. et al. **nlme**: linear and nonlinear mixed effects models. R package version 3.1-122. 2015. <URL:<http://CRAN.R-project.org/package=nlme>>. Acesso em: 7 set. 2015.

R DEVELOPMENT CORE TEAM. **R**: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2015. Disponível em: <<http://www.r-project.org>>. Acesso em: 21 ago. 2015.

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

RENA, A. B.; MAESTRI, M. Fisiologia do cafeeiro. **Informe Agropecuário**, Belo Horizonte, v. 11, p. 26-40, 1985.

SÁGIO, S. A. **Características fisiológicas e bioquímicas de duas cultivares de café de ciclos de maturação precoce e tardio**. 2009. 46 p. Dissertação (Mestrado em Fisiologia Vegetal) - Universidade Federal de Lavras, Lavras, 2009.

SANTORO, K. R. et al. Estimativas de parâmetros de curvas de crescimento de bovinos zebu, criados no Estado de Pernambuco. **Revista Brasileira de Zootecnia**, Viçosa, MG, v. 34, n. 6, p. 2262-2279, dez. 2005.

SANTOS, M. A. et al. Growth and development of gabiropa [*Campomanesia adamantium* (Cambess.) O. Berg] fruits. **African Journal of Agricultural Research**, Nairob, v. 10, n. 14, p. 1765-1772, Apr. 2015.

SILVA, A. C. et al. Características produtivas do cafeeiro arábica irrigado por pivô central na região de Lavras, MG. **Coffee Science**, Lavras, v. 6, n. 2, p. 128-136, maio/ago. 2011.

SILVA, D. F. P. et al. Curva de crescimento e padrão respiratório de frutos de genótipos de pessegueiro em região de clima subtropical. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 35, n. 2, p. 642-649, jun. 2013.

SHEEHY, J. E.; MITCHELL, P. L.; FERRER, A. B. Bi-Phasic growth patterns in rice. **Annals of Botany**, London, v. 94, n. 6, p. 811-817, Oct. 2004.

VÁSQUEZ, J. A. et al. Evaluation of non-linear equations to model different animal growths with mono and bisigmoid profiles. **Journal of Theoretical Biology**, London, v. 314, p. 95-105, Aug. 2012.

2 CAPÍTULO 2 Estimação de parâmetros de modelos duplo sigmoidais: Uma abordagem bayesiana utilizando reamostragem por importância ponderada.

RESUMO

No estudo de curvas de crescimento de frutos, o tamanho amostral é geralmente pequeno, assim a estimação dos parâmetros deve preferencialmente ser feita pela metodologia bayesiana, pois incorpora-se informações à priori, diminuindo os efeitos de se ter poucas observações. As expressões de modelos não lineares duplo sigmoidais são complexas, assim a distribuição à posteriori não possui forma conhecida e são necessários métodos de simulação estocástica para obter informações das marginais de cada parâmetro. A reamostragem por importância ponderada gera amostras destas distribuições, sendo mais simples e computacionalmente mais eficiente do que os algoritmos MCMC. O objetivo deste trabalho foi propor uma nova abordagem da técnica de reamostragem por importância ponderada para estimar parâmetros de modelos de regressão e selecionar o modelo duplo sigmoidal mais adequado para descrever o crescimento de frutos do cafeeiro. Foram utilizados as quatro combinações possíveis dos modelos não lineares Logístico e Gompertz para descrever o acúmulo de massa fresca em frutos de cafeeiro. Na reamostragem por importância ponderada foi utilizada uma aproximação da verossimilhança como distribuição candidata. Tal técnica foi eficiente pois forneceu parâmetros com interpretação prática e baixo esforço computacional, logo pode ser utilizada para estimar parâmetros de curvas de crescimento duplo sigmoidais. A qualidade de ajuste foi ótima nos quatro modelos de modo que todos ficaram muito semelhantes e qualquer um deles pode ser escolhido para descrever a curva de crescimento de frutos do cafeeiro.

Palavras-chave: Curva de crescimento; Duplo sigmoidal; Regressão não linear; Verossimilhança.

ABSTRACT

In the study of fruit growth curves, the sample size is usually small. Thereby, the estimation of the parameters should preferably be conducted by the Bayesian methodology, once it incorporates information a priori, reducing the effects of having only few observations. The expressions of nonlinear double sigmoidal models are complex. Therefore the posteriori distribution does not have a known form, and stochastic simulation methods are required in order to obtain information from each parameter. The weighted importance resampling generates samples of these distributions and are computationally simpler and more efficient than the known MCMC algorithms. The aim of this study was to propose a new approach in weighted importance resampling to estimate the parameters of regression models and select the most suitable double sigmoidal model to describe the growth of coffee fruits. A total of four possible combinations of nonlinear models Logistic and Gompertz were used to describe the fresh weight accumulation of coffee fruits. In the weighted importance resampling an approximation of likelihood was used to generate candidates samples. This technique was effective since it provided parameters with practical interpretation and low computational effort, therefore, it can be used to estimate parameters of double sigmoidal growth curves. The quality of adjustment was great in all four models. Due to the models similarities, any of them can be chosen to describe the growth curve of coffee fruits.

Keywords: Double sigmoid; Growth curve; Likelihood; Nonlinear regression.

2.1 INTRODUÇÃO

É comum nas mais diversas áreas do conhecimento situações em que o pesquisador tem a necessidade de estudar a relação entre duas ou mais variáveis quantitativas. Tal relação pode ser descrita por meio de funções matemáticas, as quais são chamadas modelos de regressão. Por serem capazes de descrever as mais diversas situações, apresentar melhor qualidade de ajuste e principalmente, apresentar interpretação prática para seus parâmetros, os modelos de regressão não lineares são muito utilizados, principalmente no estudo de curvas de crescimento (CARNEIRO et al., 2014; FERNANDES et al., 2014; PEREIRA et al., 2014).

Uma curva de crescimento é gerada por uma sequência de medidas, de determinada característica ligada ao desenvolvimento do fruto, em função do tempo, geralmente estas medidas são: massa, volume, comprimento ou diâmetro. As curvas de crescimento de frutos geralmente são caracterizadas por uma assíntota superior e pelo menos um ponto de inflexão, sendo muito bem descritas por sigmóides simples. No entanto, em frutos de caroço, o crescimento deste notavelmente se sobrepõe ao crescimento da polpa, fazendo com que suas medidas em função do tempo apresentem dois pontos de inflexão caracterizando uma curva com formato de S duplo que pode ser descrita por dois sigmóides (COOMBE, 1976), conforme a Figura 2.1.

Coombe (1976) afirma que todos os frutos com caroço e algumas bagas (groselha, morango) apresentam padrão de desenvolvimento sigmoidal duplo. Como pode ser observado na Figura 2.1, este autor divide o crescimento em três etapas (i, ii e iii). As etapas i e iii são de crescimento rápido e a ii de crescimento lento. As idades que delimitam cada uma destas etapas não são muito claras, pois cada espécie apresenta aspectos particulares de desenvolvimento e os parâmetros da curva de crescimento podem auxiliar na identificação destes pontos (CUNHA

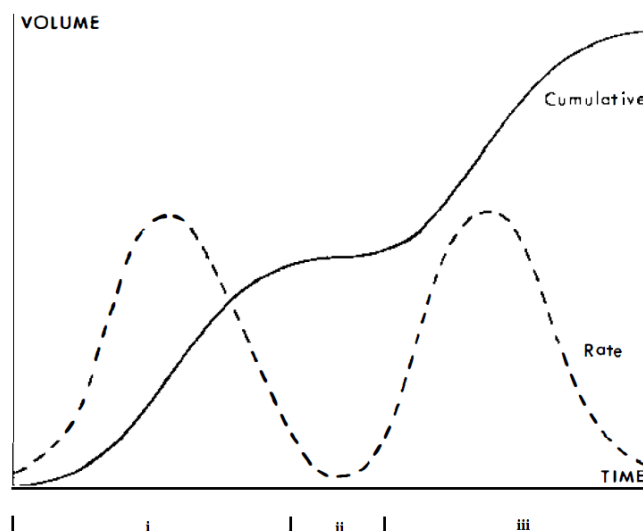


Figura 2.1 Curva de crescimento duplo sigmoidal (cumulative) e taxa de crescimento (rate) extraída de Coombe 1976.

JÚNIOR et al., 2007; NAVA et al., 2014).

As curvas duplo sigmoidais podem ser obtidas pela soma de duas funções de crescimento simples considerando-se, em uma única expressão, dois sigmóides, cada um com o objetivo de explicar uma determinada etapa do crescimento (ASCHONITIS et al., 2014; MISCHAN et al., 2015). A escolha de qual modelo deve ser colocado em cada etapa depende, obviamente, do padrão de crescimento em cada uma delas (SHEEHY; MITCHELL; FERRER, 2004).

Autores como Castro e Marraccini (2006), Cunha Júnior et al. (2007), Famiani et al. (2012), Nava et al. (2014) e Santos et al. (2015) identificaram um padrão de crescimento duplo sigmoidal em frutos de café, pêssego, ameixa, goiaba e gabioba, respectivamente. Embora consigam identificar em várias espécies este padrão de crescimento a maioria dos autores não estima os parâmetros dos modelos de regressão, talvez pelo não conhecimento sobre os modelos de regressão

não linear duplo sigmoidais, ou mesmo conhecendo, esbarram nas dificuldades envolvidas em estimar seis parâmetros para uma curva deste tipo.

A abordagem mais comum para a estimação destes parâmetros é a da inferência clássica feita geralmente pelo método de mínimos quadrados. Uma das principais dificuldades é que o sistema de equações normais (SEN) é não linear e, conseqüentemente, não possui forma fechada para as soluções, elas são obtidas por algoritmos iterativos, sendo que os mais conhecidos são: Gauss-Newton, Marquardt e Gradiente. Porém, em presença de poucos dados amostrais, quando o modelo é matematicamente complexo e/ou apresenta muitos parâmetros, a convergência pode não ser alcançada por nenhum destes algoritmos, impossibilitando o uso do modelo (GALLANT, 1987).

Além disso, na abordagem clássica o processo de estimação dos parâmetros é feito com base na teoria assintótica e, naturalmente, quando se tem poucos dados amostrais a obtenção das estimativas fica muito comprometida, gerando curvas de crescimento atípicas (MARTINS FILHO et al., 2008). No entanto, a ocorrência de poucas observações longitudinais é comum no estudo de curvas de crescimento de frutos, devido ao curto período de frutificação, impossibilitando a obtenção de muitas medidas, ao contrário do que acontece por exemplo, no crescimento animal, no qual os animais podem ser acompanhados por anos, obtendo várias mensurações.

Uma alternativa para estes problemas é a utilização da abordagem bayesiana na qual os parâmetros são considerados como variáveis aleatórias e possuem distribuição de probabilidade. Nesta metodologia são incorporadas no processo de estimação dos parâmetros informações já existentes sobre o fenômeno em estudo diminuindo assim o efeito de se ter poucos dados. A ideia está fundamentada no teorema de Bayes de modo que o conhecimento pré-existente sobre os parâmetros,

ou “à priori” ($P(\theta)$), é atualizado com base nas informações presentes nos dados por meio da “verossimilhança” ($L(Y|\theta)$) formando a distribuição “à posteriori” ($P(\theta|Y)$), representada por:

$$P(\theta|Y) \propto L(Y|\theta)P(\theta)$$

Por fim toda a inferência é feita utilizando a distribuição “a posteriori” ($P(\theta|Y)$) e as informações sobre os parâmetros podem ser resumidas, por exemplo, pela média, moda ou mediana desta distribuição (PAULINO; TURKMAN; MURTEIRA, 2003). Desejando, ainda, obter informações sobre cada parâmetro em particular integra-se a distribuição a posteriori conjunta em relação aos demais parâmetros obtendo, assim, a distribuição “marginal a posteriori” do parâmetro de interesse.

Desta forma, a integração é uma ferramenta muito importante em inferência bayesiana, tanto no cálculo de esperanças, quanto na obtenção de marginais. No entanto, exceto em casos muito simples, estas integrais não são resolvidas analiticamente. Assim, são utilizados métodos de aproximações numéricas ou técnicas de aproximação analítica.

Como comentado por Smith e Gelfand (1992), uma vez encontradas as densidades à posteriori, é possível gerar amostras das mesmas e, reciprocamente, dada uma amostra pode-se, aproximadamente, recriar a densidade. Uma das ferramentas de simulação estocástica mais conhecidas e utilizadas para gerar amostras e, conseqüentemente, obter aproximações destas integrais são os algoritmos de Monte Carlo via Cadeias de Markov (MCMC). No entanto, estes geram amostras dependentes, podem ser complicados de se implementar e, principalmente, ensinar. Além disso, os algoritmos MCMC podem ser lentos e requerem diagnósticos cuidadosos para avaliar a sua convergência, a qual inclusive

pode não ser alcançada, impossibilitando seu uso (COLE et al., 2012; STRAUB; PAPAIOANNOU, 2015).

Nesse sentido, existem algumas alternativas aos algoritmos MCMC para obter aproximações das integrais de interesse em inferência bayesiana, sendo que os principais estão fundamentadas nas técnicas de reamostragem por importância proposta por Rubin (1987). Assim como no caso do MCMC, a ideia destas técnicas é gerar amostras da distribuição à posteriori em questão, no entanto, ao contrário do MCMC, as amostras geradas são independentes, o tempo computacional para sua implementação é menor, além de possuir uma teoria mais atraente (COLE et al., 2012; LOPES; POLSON; CARVALHO, 2012).

A amostragem por importância consiste em forçar a obtenção de um maior número de amostras nas partes mais importantes do problema (regiões de maior densidade de probabilidade no espaço dos parâmetros). Esta distorção se faz introduzindo uma nova distribuição da qual seja possível gerar amostras e os valores desta devem ser corrigidos por um “fator peso” que é calculado com base na distribuição de interesse. No entanto, na proposta de Smith e Gelfand (1992) o cálculo deste fator de correção pode se tornar complicado, resultando em problemas numéricos em algumas condições, contrariando a ideia de facilitar o processo de obtenção de amostras. Assim, faz-se necessário uma alternativa para contornar esta dificuldade no cálculo dos pesos de modo a manter a metodologia de amostragem por importância ponderada atraente.

Objetivou-se neste trabalho, propor uma nova abordagem da técnica de amostragem por importância ponderada para estimar parâmetros, utilizando inferência bayesiana. Concomitantemente, selecionar o modelo de regressão não linear duplo sigmoidal mais adequado para descrever a curva de crescimento do fruto do cafeeiro.

2.2 MATERIAL E MÉTODOS

Foram utilizados dados de crescimento do fruto do cafeeiro, cultivar Obatã IAC 1669-20 da espécie *C. arabica*, extraídos de Cunha e Volpe (2011). Os autores acompanharam o desenvolvimento do fruto dos 96 dias após o florescimento (DAF) até os 293 DAF. Foram realizadas 14 coletas com um espaçamento médio entre elas de 15 dias. Os frutos eram acondicionados em sacos plásticos para posterior pesagem em balança analítica de 0,0001g de precisão e assim obter a massa fresca média em gramas (Fig. 2.2).

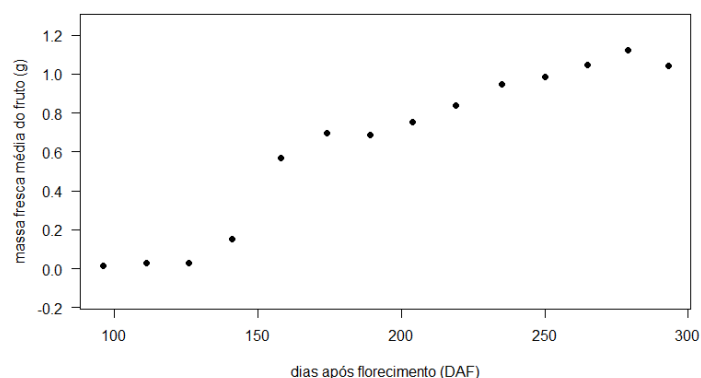


Figura 2.2 Dados de massa fresca média de frutos do cafeeiro observados por Cunha e Volpe (2011).

Para compor a expressão dos modelos de regressão duplo sigmoidais foram utilizados os modelos não lineares Logístico e Gompertz por serem os mais indicados na descrição de características que apresentem crescimento sigmoide, utilizando as parametrizações sugeridas por Fernandes et al. (2015). Conforme comentado por Sheehy, Mitchell e Ferrer (2004) a escolha do modelo a ser utilizado em cada etapa depende, naturalmente, do padrão de crescimento da mesma, assim o modelo pode ser definido como:

$$Y|\mu, \sigma^2 \sim N(\mu, I\sigma^2),$$

em que μ pode assumir as 4 combinações possíveis dos modelos Logístico (L) e Gompertz (G):

Duplo Logístico (LL):

$$\mu_{LL} = \frac{a_1}{1 + e^{k_1(b_1 - X)}} + \frac{a_2 - a_1}{1 + e^{k_2(b_2 - X)}} \quad (2.1)$$

Duplo Gompertz (GG):

$$\mu_{GG} = a_1 e^{-e^{k_1(b_1 - X)}} + (a_2 - a_1) e^{-e^{k_2(b_2 - X)}} \quad (2.2)$$

Logístico + Gompertz (LG)

$$\mu_{LG} = \frac{a_1}{1 + e^{k_1(b_1 - X)}} + (a_2 - a_1) e^{-e^{k_2(b_2 - X)}} \quad (2.3)$$

Gompertz + Logístico (GL)

$$\mu_{GL} = a_1 e^{-e^{k_1(b_1 - X)}} + \frac{a_2 - a_1}{1 + e^{k_2(b_2 - X)}} \quad (2.4)$$

em que: Y é a massa fresca do fruto; X é idade, em dias após o florescimento (DAF), nas quais foram mensuradas a massa fresca; a_1 representa a massa fresca máxima esperada para a primeira etapa de crescimento; a_2 é a massa fresca final, ou peso à maturidade do fruto; b_1 e b_2 representam a abscissa do ponto de inflexão nas respectivas etapas; k_1 e k_2 representam o índice de crescimento nas etapas 1 e 2. Quanto maior valor de k mais precoce é o fruto e menos tempo é necessário para que ele atinja seu valor máximo (a_1 ou a_2).

Dessa forma, o vetor de parâmetros para cada um dos quatro modelos é dado por: $\theta' = [a_1, b_1, k_1, a_2, b_2, k_2]$. O próximo passo para implementação de uma análise bayesiana é a especificação de distribuições “a priori”.

Elicitação das distribuições à priori

A ideia básica na elicitación de distribuições à priori para um parâmetro é, com base no conhecimento prático, estabelecer um certo intervalo de valores plausíveis e alguma medida de posição para o mesmo, com o auxílio de especialistas. A partir destes pode-se estabelecer distribuições para os parâmetros, esta técnica é conhecida como “priori de especialista” (GARTHWAITE; KADANE; O’HAGAN, 2005).

A distribuição beta é indicada para modelar uma variável aleatória V com função densidade de probabilidade definida no intervalo $(0, 1)$. Esta distribuição pode ainda ser generalizada para uma variável aleatória W que esteja definida em qualquer intervalo finito (m, n) desde que seja aplicada a seguinte relação: $W = (n - m)V + m$. Deste modo, W seguirá uma distribuição BetaGeneralizada(α, β), assim a distribuição Beta é adequada para modelar variáveis aleatórias cujos valores estejam restritos a um intervalo identificável (MCDONALD; XU, 1995).

Nesse trabalho todas as distribuições a priori utilizadas são Betas, pois conforme Moala e Penha (2016), esta é uma das distribuições mais utilizadas para modelar a priori de especialista. Assim as prioris para cada um dos seis parâmetros tem a seguinte expressão:

$$P(\theta) \propto \theta^{\alpha-1}(1 - \theta)^{\beta-1}, \quad (2.5)$$

em que α e β são hiperparâmetros (conhecidos) da distribuição Beta.

Priori para o parâmetro a_1

Conforme comentado anteriormente o parâmetro a_1 representa a massa fresca do fruto ao final da primeira etapa, ou primeiro sigmóide, do crescimento. O valor máximo de massa fresca observado por Cunha e Volpe (2011) foi 1,12 gramas. De modo natural, observando o gráfico de dispersão dos dados (Figura 2.2) pressupõe-se que a estimativa do parâmetro a_1 seja em torno da metade do máximo observado. Assim, a distribuição a priori escolhida foi uma Beta(2,2), isto é, uma distribuição Beta com hiperparâmetros $\alpha = 2$ e $\beta = 2$ cujo formato pode ser observado na Figura 2.3 (a).

Priori para o parâmetro b_1

Este parâmetro corresponde a abscissa do ponto de inflexão do primeiro sigmóide. Observando a Figura 2.2, percebe-se que esta etapa de crescimento dura, mais ou menos, até os 200 DAF. Assim, uma escolha natural para priori do parâmetro b_1 é uma distribuição $Beta(\alpha, \beta)$ definida no intervalo (96,200), pois a primeira coleta ocorreu aos 96 DAF e o primeiro sigmóide dura até os 200 DAF. Nesse caso é necessário aplicar a transformação sugerida por McDonald e Xu (1995).

Seja $V \sim Beta(\alpha, \beta)$ definida no intervalo (0,1), então a variável aleatória b_1 definida por:

$$b_1 = (200 - 96) * V + 96$$

$$b_1 = 104 * V + 96$$

tem distribuição BetaGeneralizada(α, β) definida no intervalo (96,200). Assim a distribuição a priori escolhida para o parâmetro b_1 é $[104*Beta(2,2) + 96]$, cuja densidade está na Figura 2.3 (b).

Priori para o parâmetro k_1

Este parâmetro, definido anteriormente como índice de crescimento, recebe em geral valores pequenos, concentrados próximo ao limite inferior do intervalo (0,1) na qual está definida a distribuição Beta. Assim, como distribuição a priori para k_1 foi escolhida uma Beta(2,10), conforme Figura 2.3 (c).

Priori para o parâmetro a_2

Este é o parâmetro de maior interesse, pois representa o peso final do fruto, sendo de fundamental importância uma vez que o café é comercializado em sacas de 60kg. Como comentado na definição da priori para o parâmetro a_1 , é natural pressupor que aquele parâmetro seja a metade da massa fresca final do fruto. Assim uma priori lógica para o parâmetro a_2 é o multiplicar a priori do parâmetro a_1 por 2, ficando portanto esta priori definida como $2 * \text{Beta}(2,2)$ (Figura 2.3 (d)).

Priori para o parâmetro b_2

Este parâmetro corresponde a abscissa do ponto de inflexão da segunda etapa de crescimento, portanto ele ocorre depois do b_1 . Como a priori do parâmetro b_1 foi definida em um intervalo (0,200), então uma escolha natural é que a priori para o parâmetro b_2 esteja definida em um intervalo (200,400). Novamente aplica-se a transformação sugerida por McDonald e Xu (1995).

Seja $V \sim \text{Beta}(\alpha, \beta)$ definida no intervalo (0,1), então a variável aleatória b_2 definida por:

$$b_2 = (400 - 200) * V + 200$$

$$b_2 = 200 * V + 200$$

tem distribuição BetaGeneralizada(α, β) definida no intervalo (200,400).

Desta forma, a distribuição a priori escolhida para o parâmetro b_2 foi $[200 * \text{Beta}(2,2) + 200]$, cuja densidade está na Figura 2.3 (e).

Priori para o parâmetro k_2

O parâmetro k_2 corresponde ao índice de crescimento da segunda etapa, ou segundo sigmóide. Observando na Figura 2.2 percebe-se que o ganho de massa na segunda etapa de crescimento é mais lento que na primeira, em outras palavras o fruto cresce menos, e portanto espera-se que $k_2 < k_1$. Assim utilizou-se uma priori Beta(2,20), que é ainda mais concentrada no início, com valores menores que a priori de k_1 , conforme Figura 2.3 (f).

Priori conjunta para o vetor de parâmetros θ

Desta forma, assumindo independência entre as prioris, a distribuição à priori conjunta para o vetor de parâmetros fica definida da seguinte maneira:

$$P(\theta) \propto P(a_1)P(b_1)P(k_1)P(a_2)P(b_2)P(k_2)$$

$$P(\theta) \propto [\text{Beta}(2, 2)][104 \times \text{Beta}(2, 2) + 96][\text{Beta}(2, 10)][2 \times \text{Beta}(2, 2)] \times \\ [200 \times \text{Beta}(2, 2) + 200][\text{Beta}(2, 20)]$$

$$P(\theta) \propto [a_1(1 - a_1)][104b_1(1 - b_1) + 96][k_1(1 - k_1)^9][2a_2(1 - a_2)] \times \\ [200b_2(1 - b_2) + 200][k_2(1 - k_2)^{19}]$$

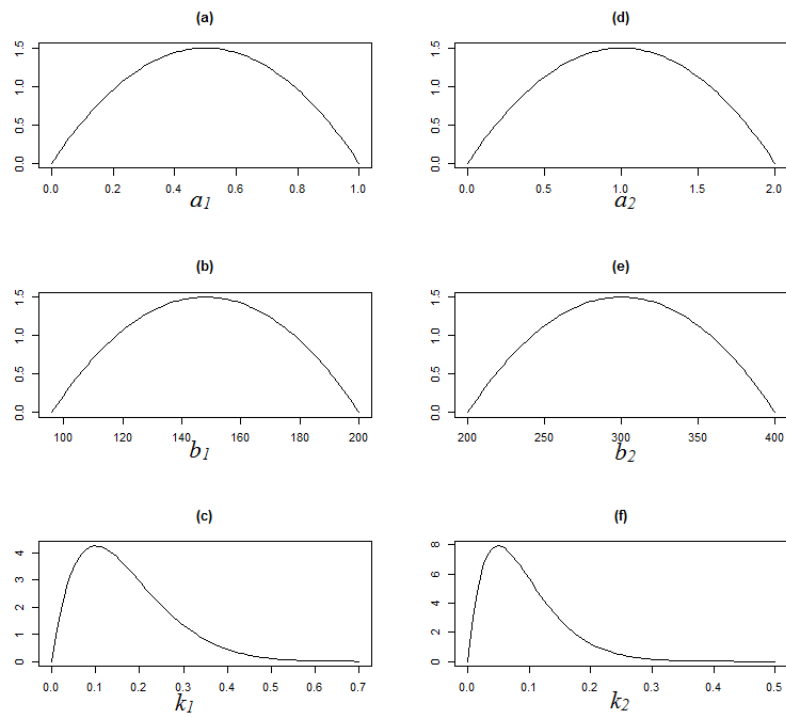


Figura 2.3 Formato das funções de densidade de probabilidade das distribuições a priori para os parâmetros a_1 (a), b_1 (b), k_1 (c), a_2 (d), b_2 (e) e k_2 (f).

Expressão da Verossimilhança $L(Y|\theta)$

Sob a suposição de normalidade do modelo, a verossimilhança é dada por:

$$L(Y|\theta) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left\{\frac{-(Y - \mu)'(Y - \mu)}{2\sigma^2}\right\}$$

De modo que conforme varia o modelo μ assume as expressões μ_{LL} , μ_{GG} , μ_{LG} e μ_{GL} .

Distribuições à posteriori $P(\theta|Y)$

Como indicado pelo teorema de Bayes a posteriori é a combinação da distribuição à priori com a verossimilhança dada por $P(\theta|Y) \propto L(Y|\theta)P(\theta)$.

Deste modo, a expressão da posteriori é a seguinte:

$$P(\theta|Y) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left\{\frac{-(Y - \mu)'(Y - \mu)}{2\sigma^2}\right\} \times P(\theta)$$

$$P(\theta|Y) \propto \exp\left\{\frac{-(Y - \mu)'(Y - \mu)}{2\sigma^2}\right\} \times P(\theta)$$

A expressão apresentada acima é a distribuição à posteriori conjunta dos modelos. Para obter informações sobre o comportamento a posteriori de cada parâmetro, média e intervalo de credibilidade por exemplo, é necessário o cálculo de integrais do tipo:

$$E[h(\theta)|Y] = \int_D h(\theta)P(\theta|Y)d(\theta)$$

em que $h(\theta)$ é uma função convenientemente escolhida.

No entanto, estas integrais são impossíveis de serem resolvidas analiticamente e será utilizada uma nova abordagem da técnica de reamostragem

por importância ponderada apresentada em Smith e Gelfand (1992) para gerar amostras das distribuições de interesse e assim obter uma aproximação da mesma.

A metodologia proposta em Smith e Gelfand (1992) consiste na seguinte ideia. Suponha que seja necessário gerar amostras de uma distribuição $f(\boldsymbol{\theta})$, mas que tal distribuição não possui forma conhecida dificultando a geração de amostras diretamente dela. Considere então uma distribuição $g(\boldsymbol{\theta})$ que seja conhecida e possível de gerar amostras definida no mesmo domínio (ou espaço paramétrico) de $f(\cdot)$. Baseado na teoria de amostragem por importância (RUBIN, 1987) multiplique e divida a expressão de $f(\cdot)$ pela expressão $g(\cdot)$, obtendo:

$$\frac{f(\boldsymbol{\theta})}{g(\boldsymbol{\theta})} \times g(\boldsymbol{\theta})$$

Agora você irá gerar amostras de $g(\boldsymbol{\theta})$ e ponderar por pesos calculados com base em $\frac{f(\boldsymbol{\theta})}{g(\boldsymbol{\theta})}$. No fim tem-se uma amostra aproximada de $f(\boldsymbol{\theta})$. O algoritmo se desenvolve como o seguinte:

i) gere uma amostra de tamanho M , $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_M)$ da distribuição $g(\boldsymbol{\theta})$;

ii) calcule os pesos para cada amostra gerada:

$$w_j = \frac{f(\boldsymbol{\theta}_j)}{g(\boldsymbol{\theta}_j)}$$

em que $1 \leq j \leq M$;

iii) normalize estes pesos:

$$q_j = \frac{w_j}{\sum_t w_t}$$

iv) retire uma amostra de tamanho m tal que $(m \leq M)$ da amostra $(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_M)$ pegando apenas os $\boldsymbol{\theta}_j$ com os maiores valores de q_j , isto é, faça uma

reamostragem ponderada por pesos calculados com base na distribuição de interesse, obtendo uma nova amostra $(\theta_1^*, \theta_2^*, \dots, \theta_m^*)$.

Smith e Gelfand (1992) mostraram que a amostra obtida no passo *iv*) forma uma amostra de uma distribuição $f(\theta^*)$ que converge em distribuição para $f(\theta)$ conforme m cresce. Perceba que a única exigência para a distribuição geradora de candidatos $g(\theta)$ é que esta esteja definida no mesmo espaço paramétrico que a distribuição de interesse $f(\theta)$.

Em problemas de inferência bayesiana, $f(\theta)$ é a distribuição à posteriori ($P(\theta|Y)$) e uma escolha natural proposta por Smith e Gelfand (1992) é considerar $g(\theta)$ como a distribuição à priori. Assim, inicialmente gera-se uma amostra de tamanho M da distribuição à priori e o cálculo dos pesos no passo *ii*) fica:

$$w_j = \frac{f(\theta_j)}{g(\theta_j)} = \frac{P(\theta_j|Y)}{P(\theta_j)} = \frac{L(Y|\theta_j)P(\theta_j)}{P(\theta_j)} = L(Y|\theta_j)$$

e no passo *iii*):

$$q_j = \frac{w_j}{\sum_t^M w_t} = \frac{L(Y|\theta_j)}{\sum_t^M L(Y|\theta_t)}$$

Observe que neste último passo é necessário calcular a verossimilhança para todos os pontos amostrados e tal cálculo pode se tornar trabalhoso, em certas situações, contradizendo a ideia de simplicidade que é a grande vantagem do método de amostragem por importância em relação aos algoritmos MCMC, por exemplo. Pode-se ainda cair em problemas numéricos se a geradora de candidatos for muito ruim, levando a uma grande variabilidade dos pesos. Nesse sentido, o presente trabalho propõe que seja considerada a distribuição $g(\theta)$ como uma aproximação da verossimilhança, denotada por $Q(\theta)$. A distribuição geradora de candidatos $Q(\theta)$ é uma aproximação da verossimilhança real, criada com base no

ponto de máximo da verossimilhança e na decomposição de Cholesky da matriz de covariâncias dos parâmetros. Como assume-se distribuição normal para o modelo, a medida que $M \rightarrow +\infty$, tem-se que $Q(\theta)$ converge em distribuição para $L(Y|\theta)$.

Assim, inicialmente toma-se uma amostra de tamanho M da verossimilhança e os pesos são calculados da seguinte maneira:

$$w_j = \frac{f(\theta_j)}{g(\theta_j)} = \frac{P(\theta_j|Y)}{Q(\theta_j)} = \frac{L(Y|\theta_j)P(\theta_j)}{Q(\theta_j)} \simeq \frac{L(Y|\theta_j)P(\theta_j)}{L(Y|\theta_j)} = P(\theta_j)$$

e no passo *iii*):

$$q_j = \frac{w_j}{\sum_t^M w_t} = \frac{P(\theta_j)}{\sum_t^M P(\theta_t)} = P(\theta_j). \quad (2.6)$$

O denominador da expressão acima, tende para 1 a medida em que M cresce, logo os pesos são obtidos diretamente na distribuição à priori, o que facilita os cálculos.

Em outras palavras, a metodologia aqui proposta sugere considerar uma aproximação da verossimilhança como distribuição candidata $g(\theta)$, o que simplifica o cálculo dos pesos para a reamostragem uma vez que estes podem ser obtidos diretamente da distribuição à priori, sem precisar calcular a verossimilhança em todos os pontos amostrados. Deste modo a ideia de amostragem por importância ponderada fica, ainda mais, atraente e fácil de ser compreendida e, principalmente ensinada.

Na comparação dos modelos foi utilizado o fator de Bayes que, de modo geral, corresponde a uma razão de chances a posteriori, seguindo um processo análogo ao do teste de razão de verossimilhanças utilizado na inferência clássica.

Sua expressão é dada por:

$$FB_{(M_1, M_2)} = \frac{\hat{P}(Y|M_1)}{\hat{P}(Y|M_2)} = \frac{\frac{1}{m} \sum_{q=1}^m L(Y|\theta_q, M_1)}{\frac{1}{m} \sum_{q=1}^m L(Y|\theta_q, M_2)},$$

em que: $\hat{P}(Y|M_1)$ e $\hat{P}(Y|M_2)$ são a verossimilhança marginal estimada para os modelos 1 e 2 respectivamente; e o índice q representa cada valor de θ_q amostrado tal que $1 \leq q \leq m$.

A interpretação do fator de Bayes é simples e direta: caso $FB_{(M_1, M_2)} > 1$ tem-se que o modelo do numerador é mais plausível que o do denominador, caso contrário, se $FB_{(M_1, M_2)} < 1$ então o modelo do denominador é mais plausível que o do numerador.

A média e a variância da verossimilhança marginal à posteriori do modelo também foram calculadas de modo a dar uma ideia sobre seu comportamento. Dessa forma, conforme sugerido por Raftery et al. (2007) calculou-se também a aproximação de Monte Carlo do Critério de Informação de Akaike (AIC). Sendo este obtido por:

$$AICM = 2(\bar{\ell} - S_{\bar{\ell}}^2),$$

em que: AICM é o Critério de Informação de Akaike de Monte Carlo; $\bar{\ell}$ é média amostral e $S_{\bar{\ell}}^2$ é a variância amostral da verossimilhança marginal do modelo. Quanto maior o valor de AICM melhor é o modelo (RAFTERY et al., 2007).

Foram obtidos também os intervalos de credibilidade de máxima densidade a posteriori (do inglês HPD) com o nível de credibilidade fixado em 95%.

Na análise dos dados, processos de amostragem e reamostragem para obtenção das estimativas para os parâmetros, bem como toda a parte computacional envolvida, foi utilizado o software estatístico R (R CORE TEAM, 2015). Na

análise de diagnóstico e resumo das cadeias utilizou-se o pacote *coda* do R.

2.3 RESULTADOS E DISCUSSÃO

Como apresentado na metodologia, a distribuição $f(\boldsymbol{\theta})$ da qual deseja-se gerar amostras, neste trabalho é: $P_{LL}(\boldsymbol{\theta}|Y)$, $P_{GG}(\boldsymbol{\theta}|Y)$, $P_{LG}(\boldsymbol{\theta}|Y)$ e $P_{GL}(\boldsymbol{\theta}|Y)$. Estas são as distribuições à posteriori conjunta de cada um dos quatro modelos e deseja-se obter amostras das marginais para cada parâmetro do vetor $\boldsymbol{\theta}' = [a_1, b_1, k_1, a_2, b_2, k_2]$.

A geradora de candidatos escolhida foi a aproximação da verossimilhança ($Q(\boldsymbol{\theta})$). Para estimar $Q(\boldsymbol{\theta})$, o ponto de máximo da verossimilhança ($\hat{\boldsymbol{\theta}}$) bem como a matriz de covariâncias (Σ) dos parâmetros do respectivo modelo, foram obtidos utilizando a função *gnls* presente no pacote *nlme* do software R. Deste modo o algoritmo de obtenção das amostras foi implementado para todos os quatro modelos duplo sigmoidais como se segue:

- i) gerou-se uma amostra de tamanho $M=10000$ de $Q(\boldsymbol{\theta})$ do respectivo modelo;

Para gerar valores de $Q(\boldsymbol{\theta})_j$ assumiu-se que $Q(\boldsymbol{\theta}) \sim N_6(\hat{\boldsymbol{\theta}}, \Sigma)$.

Depois foi feita a decomposição de Cholesky (L) de Σ , tal que $\Sigma = LL'$ e simulado um vetor Z_j que tem distribuição $N_6(\emptyset, I)$.

Utilizando a transformação linear $Q(\boldsymbol{\theta})_j = \hat{\boldsymbol{\theta}} + LZ_j$ obteve-se o vetor desejado. Este processo foi repetido dez mil vezes.

- ii) calculou-se os pesos, na distribuição a priori $P(\boldsymbol{\theta})$, para cada valor gerado;
- iii) na normalização dos pesos eles permanecem os mesmos, uma vez que o denominador é igual a 1 (expressão 2.6);

- iv) foi feita uma reamostragem, com reposição, ponderada pelos pesos obtidos da distribuição à priori de tamanho $m=10000$, na amostra obtida no primeiro passo.

Conforme mostrado por Smith e Gelfand (1992) a amostra obtida no quarto passo converge em distribuição para a distribuição à posteriori do respectivo modelo, e informações marginais sobre cada um destes parâmetros são agora facilmente obtidas. Este processo de reamostragem ponderada foi repetido para cada um dos quatro modelos estudados.

Deste modo foram aproximadamente recriadas, a partir de amostras, a densidade marginal de cada parâmetro. Visando obter uma estimativa pontual, para fins de substituir na expressão do modelo e assim obter a curva estimada, calculou-se as médias amostrais e o erro padrão da média em cada distribuição marginal dos parâmetros. Deste modo, são apresentados nas Tabelas 2.1 à 2.4 as médias da distribuição marginal à posteriori, o desvio padrão e o intervalo HPD (95%) para os parâmetros dos modelos duplo sigmoidais.

As Figuras 2.4 à 2.7 apresentam a densidade das distribuições marginais para os parâmetros destes modelos. Nestas figuras fica evidenciada uma vantagem prática da utilização da metodologia bayesiana, pois são apresentadas as distribuições de probabilidade marginais dos parâmetros do modelo (MARTINS FILHO et al., 2008). E assim o pesquisador tem muito mais informações para concluir diretamente sobre o comportamento destes parâmetros.

Tabela 2.1 Média, desvio padrão e intervalo HPD para os parâmetros do modelo Duplo Logístico (LL).

Parâmetro	média	Desvio padrão	HPD 95%	
			LI	LS
a_1	0,6708	0,0421	0,5863	0,7494
b_1	148,1257	1,5725	145,0599	151,1921
k_1	0,1659	0,0284	0,1099	0,2194
a_2	1,0817	0,0258	1,0316	1,1336
b_2	225,6010	5,1857	215,1511	235,7878
k_2	0,0631	0,0191	0,0239	0,0998

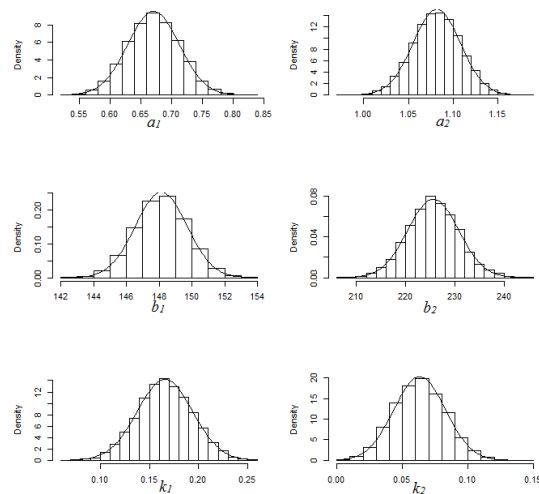


Figura 2.4 Funções de densidade de probabilidade das distribuições marginais de cada parâmetro do modelo duplo Logístico.

Tabela 2.2 Média, desvio padrão e intervalo HPD para os parâmetros do modelo Duplo Gompertz (GG).

Parâmetro	média	Desvio padrão	HPD 95%	
			LI	LS
a_1	0,7062	0,0311	0,6436	0,7645
b_1	144,6295	1,2017	142,2985	146,9263
k_1	0,1163	0,0215	0,0731	0,1569
a_2	1,0868	0,0312	1,0248	1,1482
b_2	220,3388	4,9597	210,5473	230,2800
k_2	0,0491	0,0144	0,0209	0,0782

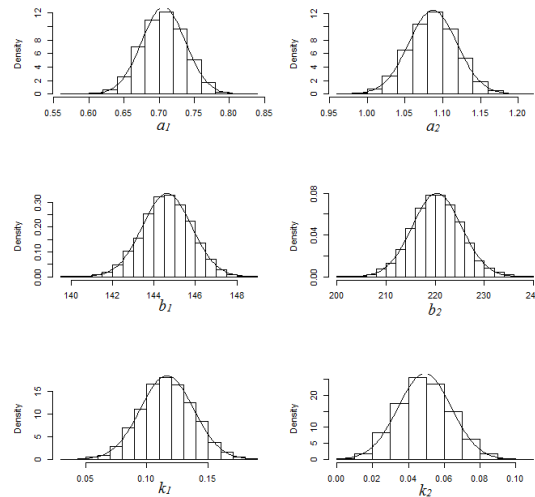


Figura 2.5 Funções de densidade de probabilidade das distribuições marginais de cada parâmetro do modelo duplo Gompertz.

O parâmetro a_2 , que é o de maior interesse por se tratar do peso final do fruto, sendo o café comercializado em sacas de acordo com seu peso, foi muito bem estimado, de modo que todos os intervalos de credibilidade para a_2 contém as últimas observações de massa fresca relatadas por Cunha e Volpe (2011).

Tabela 2.3 Média, desvio padrão e intervalo HPD para os parâmetros do modelo Logístico + Gompertz (LG).

Parâmetro	média	Desvio padrão	HPD 95%	
			LI	LS
a_1	0,6967	0,0269	0,6444	0,7492
b_1	148,6831	1,4223	145,9068	151,3020
k_1	0,1582	0,0234	0,1122	0,2019
a_2	1,0896	0,0304	1,0303	1,1488
b_2	219,3103	4,4855	210,6150	228,3926
k_2	0,0472	0,0128	0,0221	0,0722

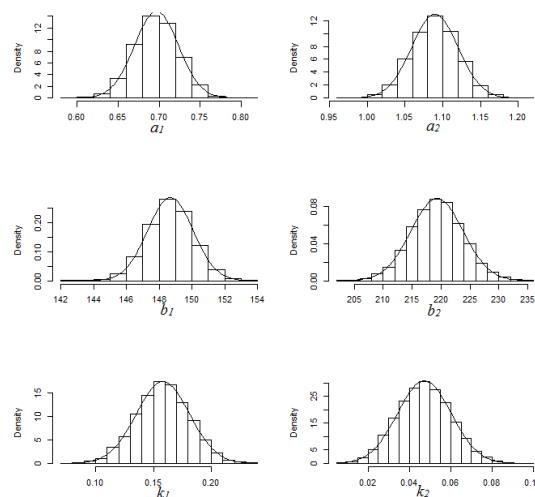


Figura 2.6 Funções de densidade de probabilidade das distribuições marginais de cada parâmetro do modelo Logístico + Gompertz.

As principais diferenças nas estimativas são observadas nos parâmetros b_1 e b_2 , que indicam a abscissa do ponto de inflexão, pois assim como comentado por Fernandes et al. (2015) o modelo de Gompertz é mais precoce e tende a atingir tal ponto antes do que o Logístico. Deste modo, as estimativas para os parâmetros b_1 e b_2 tendem a ser menores quando o modelo em questão é o de Gompertz. Por

exemplo, a estimativa de b_1 na Tabela 2.4 é menor que a estimativa de b_1 na Tabela 2.3, pois embora estimem a mesma característica (abscissa do ponto de inflexão da primeira etapa de crescimento), na Tabela 2.4 foi utilizado o modelo de Gompertz para modelar a primeira etapa de crescimento e este tende a atingir o ponto de inflexão antes do que o modelo Logístico, que foi utilizado na Tabela 2.3.

Tabela 2.4 Média, desvio padrão e intervalo HPD para os parâmetros do modelo Gompertz + Logístico (GL).

Parâmetro	média	Desvio padrão	HPD 95%	
			LI	LS
a_1	0,5722	0,0755	0,4283	0,7156
b_1	143,6264	1,3797	140,9495	146,1448
k_1	0,1584	0,0615	0,0452	0,2779
a_2	1,1227	0,0463	1,0286	1,2140
b_2	218,5657	9,8427	199,4881	238,4922
k_2	0,0368	0,0085	0,0204	0,0531

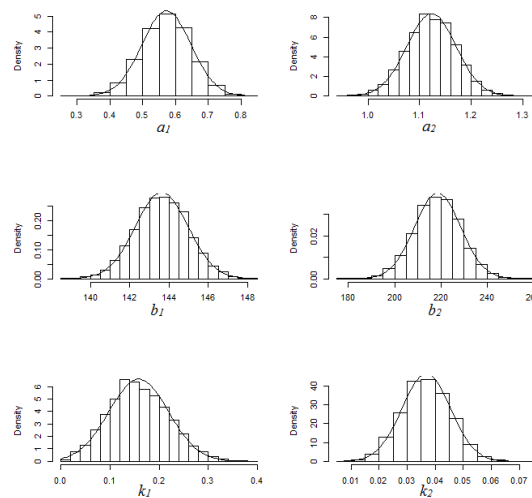


Figura 2.7 Funções de densidade de probabilidade das distribuições marginais de cada parâmetro do modelo Gompertz + Logístico.

Percebe-se nas Tabelas 2.1, 2.2, 2.3 e 2.4 que as estimativas do parâmetro k_1 são sempre maiores que as do parâmetro k_2 , sugerindo que o desenvolvimento do fruto do café é mais acelerado na primeira etapa de crescimento (ou primeiro sigmóide). Esta rápida expansão na primeira etapa de desenvolvimento do fruto do café pode ser explicada pela escala fenológica proposta por Nunes et al. (2010), na qual o estágio de crescimento rápido do fruto coincide com esta primeira etapa. Comportamento diferente do que acontece com o fruto do pêssego, por exemplo, que apresenta um crescimento mais notável na segunda etapa como comentado por Silva et al. (2013).

Outro aspecto interessante a ser comentado é a modificação nas distribuições de densidade quando comparadas com as densidades utilizadas à priori (Figura 2.3). Percebe-se que as distribuições marginais à posteriori (Figuras 2.4, 2.5, 2.6 e 2.7) ficaram bem mais concentradas, com uma menor variância, tal transformação foi causada pela incorporação das informações presente nos dados.

Pode ser observado também um desvio padrão das estimativas muito baixo, fato que indica a alta confiabilidade nos parâmetros estimados pelo modelo. Justificando assim a eficiência da metodologia bayesiana na estimação de parâmetros de modelos de regressão não linear, como também foi observado por autores como Blasco et al. (2003) e Martins Filho et al. (2008).

A obtenção das amostras, o cálculo dos pesos e a reamostragem ponderada foi feita rapidamente (cerca de 2 segundos para cada modelo) sem necessitar de muito esforço computacional. Utilizando reamostragem por rejeição, que é um pouco diferente da amostragem ponderada, mas também é baseado na teoria de reamostragem por importância, Cole et al. (2012) concluíram que o tempo computacional é muito menor se comparado com o necessário para fazer a implementação do mesmo problema via algoritmos MCMC.

Como comentado por Straub e Papaionou (2015), uma das grandes vantagens do método de reamostragem é que não há necessidade de realizar análise de convergência das cadeias, uma vez que estas foram obtidas por cálculos exatos, e não aproximados como no caso dos algoritmos MCMC. Assim este método além de ganhar em eficiência computacional, evita os possíveis problemas no diagnóstico de convergência das cadeias de Markov (LOPES; POLSON; CARVALHO, 2012).

Para comparar a qualidade de ajuste fornecida pelo modelo foi calculado o Fator de Bayes, cujos resultados estão apresentados na Tabela 2.5. Nesta tabela os modelos apresentados na linha foram sempre colocados no numerador e os da coluna no denominador.

Tabela 2.5 Fator de Bayes para a comparação dos modelo de regressão não linear utilizados.

linha/coluna	LL	GG	LG	GL
LL	1	1,0088	0,9899	1,0162
GG		1	0,9851	1,0073
LG			1	1,0163
GL				1

Os valores obtidos pelo fator de bayes não elegeram um melhor modelo, pois em todas as seis combinações possíveis o fator de bayes ficou muito próximo de 1, tornando-se inconclusivo. De acordo com a tabela proposta por Jeffreys (1961), se o resultado do fator de bayes for menor que 3,2 não existem evidências suficientes a favor do modelo do numerador. Assim, utilizando este critério de qualidade não é possível indicar um modelo com desempenho superior aos demais.

A média e a variância da verossimilhança marginal de um modelo podem ajudar na seleção de modelos, afinal aquele que apresentar a maior média e a menor variância tende a ser melhor, pois apresentam valores amostrais mais plausíveis

e com menor variabilidade. Com base nas estimativas amostrais de média e variância da verossimilhança marginal dos modelos também foi possível estimar o Critério de Informação de Akaike de Monte Carlo - AICM, como apresentado por Raftery et al. (2007).

Tabela 2.6 Média ($\bar{\ell}$) e variância (S_{ℓ}^2) da verossimilhança marginal do modelo, bem como o valor estimado do critério de informação de Akaike de Monte Carlo para os quatro modelos duplo sigmoidais.

Modelos	$\bar{\ell}$	S_{ℓ}^2	AICM
LL	0,002152	$7,39 \times 10^{-6}$	0,00429
GG	0,002133	$7,06 \times 10^{-6}$	0,00425
LG	0,002152	$7,06 \times 10^{-6}$	0,00429
GL	0,002117	$9,16 \times 10^{-6}$	0,00422

Novamente nos resultados apresentados na Tabela 2.6, não existe uma preferência muito clara por algum modelo, estando os resultados muito próximos. Por exemplo, o modelo Gompertz + Logístico possui a menor média da verossimilhança, no entanto, possui também o menor valor de AICM e a maior variância da verossimilhança. Tal fato pode ser observado também na Figura 2.8, pois o intervalo de 95% de credibilidade para este modelo fica mais amplo devido a maior variabilidade. A qualidade dos ajustes dos modelos deve realmente ser parecida, pois todos forneceram excelentes ajustes e ficaram muito semelhantes, conforme pode ser observado na Figura 2.8.

Os quatro modelos duplo sigmoidais estudados descreveram muito bem a curva de crescimento do fruto do cafeeiro, conforme Figura 2.8. Os intervalos de credibilidade obtidos utilizando a distribuição preditiva ficaram bem estreitos, dificultando assim a eleição do melhor modelo, uma vez que todos forneceram ótimos ajustes (Figura 2.8 e Tabelas 2.5 e 2.6). Confirmando, portanto, que o crescimento do fruto do cafeeiro é duplo sigmoidal, fato que corrobora com o que foi observado por autores como Castro e Marrachini (2006) e Laviola et al. (2007),

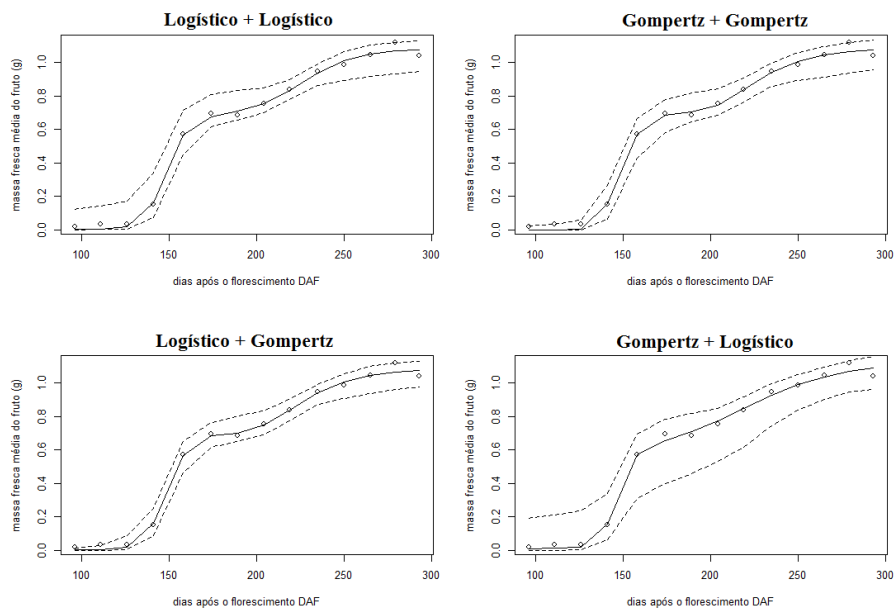


Figura 2.8 Ajuste e intervalo de credibilidade fornecido pelos 4 modelos, utilizando como estimativas dos parâmetros as médias das distribuições à posteriori marginais.

dentre outros.

A qualidade de ajuste obtida foi muito boa e os parâmetros estimados estão condizentes com a literatura, reforçando a eficiência desta metodologia de amostragem por importância como alternativa para a estimação de parâmetros de modelos de regressão duplo sigmoidais. Assim, tal abordagem pode ser utilizada para estimar os parâmetros das curvas de crescimento de outros frutos como pêsego, ameixa, goiaba e gabioba que, conforme identificado por Cunha Júnior et al. (2007), Famiani et al. (2012), Nava et al. (2014) e Santos et al. (2015), também apresentam padrão de crescimento duplo sigmoidal.

2.4 CONCLUSÕES

A modificação proposta na metodologia de reamostragem por importância mostrou-se eficiente na estimação de parâmetros dos modelos de regressão não lineares duplo sigmoidais, fornecendo parâmetros com interpretação prática e consistentes com a literatura.

As quatro combinações possíveis dos modelos Logístico e Gompertz para compor uma expressão duplo sigmoideal apresentaram ótima qualidade de ajuste, assim, qualquer uma delas pode ser utilizada na descrição da curva de crescimento de frutos do cafeeiro.

REFERÊNCIAS

ASCHONITIS, V. G. et al. A review and synthesis of bivariate non-linear models to describe the relative variation of ecological, biological and environmental parameters. **Environmental Modelig and Assessment**, London, v. 20, p. 169-182, July 2014.

BLASCO, A. et al. Bayesian analysis of the effect of selection for growth rate on growth curves in rabbits. **Genetics Selection Evolution**, Les Ulis, v. 35, n. 1, p. 21-41, Jan. 2003.

CARNEIRO, A. P. S. et al. Identidade de modelos não lineares para comparar curvas de crescimento de bovinos da raça Tabapuã. **Pesquisa Agropecuária Brasileira**, Brasília, v. 49, n. 1, p. 57-62, jan. 2014.

CASTRO, R. D.; MARRACCINI, P. Cytology, biochemistry and molecular changes during coffee fruit development. **Brazilian Journal of Plant Physiology**, Campos dos Goytacazes, v. 18, n. 1, p. 175-199, Jan./Mar. 2006.

COLE, S. R. et al. Bayesian posterior distributions without markov chains. **American Journal of Epidemiology**, Oxford, v. 175, n. 5, p. 368-375, Feb. 2012.

COOMBE, B. G. The development of fleshy fruits. **Annual Review of Plant Physiology**, Palo Alto, v. 27, p. 507-528, 1976.

CUNHA, A. R.; VOLPE, C. A. Curvas de crescimento do fruto de cafeeiro cv. Obatã IAC 1669-20 em diferentes alinhamentos de plantio. **Semina: Ciências Agrárias**, Londrina, v. 32, n. 1, p. 49-62, jan./mar. 2011.

CUNHA JUNIOR, L. C. et al. Caracterização da curva de maturação de pêsegos “Aurora-1”, na região de Jaboticabal-SP. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 29, n. 3, p. 661-665, dez. 2007.

FAMIANI, F. et al. Development and metabolism of the fruit and seed of the Japanese plum Ozark premier (Rosaceae). **Journal of Plant Physiology**, Stuttgart, v. 169, p. 551-560, Apr. 2012.

FERNANDES, T. J. et al. Parameterization effects in nonlinear models to describe growth curves. **Acta Scientiarum. Technology**, Maringá, v. 37, n. 4, p. 397-402, Oct./Dec. 2015.

FERNANDES, T. J. et al. Seleção de modelos não lineares para descrição das curvas de crescimento do fruto do cafeeiro. **Coffee Science**, Lavras, v. 9, n. 2, p. 207-215, abr./jun. 2014.

GALLANT, A. R. **Nonlinear statistical models**. New York: J. Wiley, 1987. 610 p.

GARTHWAITE, P. H.; KADANE, J. B.; O'HAGAN, A. Statistical methods for eliciting probability distributions. **Journal of the American Statistical Association**, New York, v. 100, n. 470, p. 680-701, June 2005.

JEFFREYS, H. **Theory of Probability**. London: Oxford University Press, 1961. 325 p.

LAVIOLA, B. G. et al. Acúmulo de nutrientes em frutos de cafeeiro em quatro altitudes de cultivo: cálcio, magnésio e enxofre. **Revista Brasileira de Ciência do Solo**, Campinas, v. 31, p. 1451-1462, maio 2007.

LOPES, H. F.; POLSON, N. G.; CARVALHO, C. M. Bayesian statistics with a smile: a resampling-sampling perspective. **Brazilian Journal of Probability and Statistics**, São Paulo, v. 26, n. 4, p. 358-371, 2012.

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.

MCDONALD, J. B.; XU, Y. J. A generalization of the beta distribution with applications. **Journal of Econometrics**, Amsterdam, v. 66, n. 1, p. 133-152, Mar./Apr. 1995.

MISCHAN, M. M. et al. Inflection and stability points of diphasic logistic analysis of growth. **Scientia Agricola**, Piracicaba, v. 72, n. 3, p. 215-220, May/June 2015.

MOALA, F. A.; PENHA, D. L. Métodos de elicitação da distribuição à priori Beta. **Revista Brasileira de Biometria**, São Paulo, v. 34, n. 1, p. 49-62, jan. 2016.

NAVA, A. D. et al. Growth kinetics of vegetative and reproductive organs of guava (*Psidium guajava* L.) in Iguala Guerrero, Mexico. **Agricultural Sciences**, Cambridge, v. 5, p. 1468-1475, Dec. 2014.

NUNES, F. L. et al. Modelos agrometeorológicos de estimativa da duração do estágio floração-maturação para três cultivares de café arábica. **Bragantia**, Campinas, v. 69, n. 4, p. 1011-1018, dez. 2010.

PAULINO, C. D.; TURKMAN, M. A. A.; MURTEIRA, B. **Estatística Bayesiana**. Lisboa: Fundação Calouste Gulbenkian, 2003. 446 p.

PEREIRA, A. A. et al. Descrição do crescimento vegetativo do cafeeiro cultivar Rubi MG 1192, utilizando modelos de regressão. **Coffee Science**, Lavras, v. 9, n. 2, p. 236-271, abr./jun. 2014.

R DEVELOPMENT CORE TEAM. **R**: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2015. Disponível em: <<http://www.r-project.org>>. Acesso em: 21 ago. 2015.

RAFTERY, A. E. et al. Estimating the integrated likelihood via posterior simulation using the harmonic mean identity. **Bayesian Statistics**, Kyoto, v. 8, p. 1-45, 2007.

RUBIN, D. B. A noniterative sampling/importance resampling alternative to the data augmentation algorithm for creating a few imputations when fractions of missing information are modest: the SIR algorithm. **Journal of the American Statistical Association**, New York, v. 82, n. 398, p. 543-546, 1987.

SANTOS, M. A. et al. Growth and development of gabirola [*Campomanesia adamantium* (Cambess.) O. Berg] fruits. **African Journal of Agricultural Research**, Nairobi, v. 10, n. 14, p. 1765-1772, Apr. 2015.

SILVA, D. F. P. et al. Curva de crescimento e padrão respiratório de frutos de genótipos de pessegueiro em região de clima subtropical. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 35, n. 2, p. 642-649, jun. 2013.

SHEEHY, J. E.; MITCHELL, P. L.; FERRER, A. B. Bi-phasic growth patterns in rice. **Annals of Botany**, London, v. 94, n. 6, p. 811-817, Oct. 2004.

SMITH, A. F. M.; GELFAND, A. E. Bayesian statistics without tears: a sampling-resampling perspective. **The American Statistician**, Washington, v. 46, n. 2, p. 84-88, May 1992.

STRAUB, D.; PAPAIOANNOU, I. Bayesian updating with structural reliability methods. **Journal of Engineering Mechanics**, Houston, v. 141, n. 34, p. 1-13, Mar. 2015.

3 CAPÍTULO 3 Avaliação do efeito dos alinhamentos de plantio na curva de crescimento do fruto do cafeeiro.

RESUMO

A incidência da radiação solar nas plantas do cafeeiro deve ser considerada no estudo sobre curvas de crescimento dos frutos, uma vez que interfere na fisiologia e na produtividade dos mesmos. Assim o alinhamento de plantio apresenta influência direta na fase reprodutiva do cafeeiro, de modo que o ideal seria aquele alinhamento com uniformidade na distribuição solar em ambas as faces da planta. O objetivo deste trabalho foi modelar a curva de crescimento dos frutos do cafeeiro em três alinhamentos de plantio e nas diferentes faces da planta (Norte e Sul) comparando o efeito destes fatores nos parâmetros estimados. Foram analisados dados de acúmulo de massa fresca em frutos coletados nas faces Norte e Sul de três alinhamentos de plantio, extraídos de Cunha e Volpe (2011). Foi utilizado o modelo não linear duplo Logístico e a estimação dos parâmetros feita pela metodologia bayesiana, o que possibilitou a comparação direta destes nas diferentes curvas de crescimento, conforme proposto por Silva et al. (2005). Identificou-se diferença entre as taxas de crescimento nas duas etapas do desenvolvimento do fruto, independente da face. Embora seja comprovada diferença de produtividade e na qualidade da bebida, não houve diferença entre as curvas de crescimento provenientes dos diferentes alinhamentos de plantio estudados.

Palavras-chave: Não linear; Peso à maturidade; Radiação solar; Regressão.

ABSTRACT

It is important to consider the incidence of solar radiation on coffee plants when studying the growth curves of its fruits. Solar radiation interferes with the physiology and productivity of the fruits. Thereby, planting alignment has direct influence on the reproductive phase of the coffee plant. It is important to have a uniform solar distribution in both sides of the plants. This study aims at modeling the growth curve of coffee fruits growing in three different planting alignments and in different sides of the plant (North and South), comparing the effect of these factors on the estimated parameters. Data of fresh weight accumulation in fruits collected in the north and south faces of the three planting alignments, from Cunha and Volpe (2011), was analyzed. The non-linear double Logistic model and the estimation of the parameters using the bayesian approach were used, allowing the direct comparison in the different growth curves, as proposed by Silva et al. (2005). A difference between the growth rates in the two stages of fruit development was identified, regardless the side. Although it has been proven differences in productivity and quality of coffee, there was no difference between the growth curves in the different planting alignments herein studied.

Keywords: Maturity weight; Nonlinear; Regression; Solar radiation.

3.1 INTRODUÇÃO

O café é uma das mais importantes commodities do setor agrícola brasileiro e mundial. Nesse sentido, a preocupação com a melhoria da qualidade e da produtividade leva os pesquisadores a investir cada vez mais em processos que otimizem o manejo do cafeeiro (SILVA et al., 2011). Conhecer o padrão de crescimento dos frutos pode auxiliar o produtor no desenvolvimento da cultura, indicando por exemplo, o ponto ótimo da colheita que é fundamental na obtenção de melhor qualidade da bebida (FERNANDES et al., 2014).

No estudo das curvas de crescimento do fruto do cafeeiro, um aspecto que deve ser considerado é a incidência de radiação solar nas plantas, pois ela interfere na fisiologia, produção e maturação dos frutos (CUNHA; VOLPE, 2011; CUSTÓDIO et al., 2015; OLIVEIRA et al., 2012). Nesse sentido, a distribuição espacial da lavoura por meio dos diferentes alinhamentos de plantio apresenta influência direta nas fases reprodutivas do cafeeiro e na qualidade do fruto (BARBOSA et al., 2010; CASTRO; MARRACCINI, 2006).

Trabalhos recentes revelam que a fisiologia e o desenvolvimento da semente do café variam dependendo da posição e da luminosidade que o fruto recebe, sugerindo que podem haver diferenças na bioquímica dos frutos de uma mesma planta (BARBOSA et al., 2010; CHAVES et al., 2012; CUSTÓDIO et al., 2015). A quantidade de radiação solar interceptada por uma cultura é uma importante variável meteorológica que determina o crescimento e o desenvolvimento dos frutos. Dentre os sistemas de produção, a orientação das linhas de plantio pode ser um componente bastante relevante quanto à interceptação da radiação solar direta em ambas as faces (CUNHA; VOLPE, 2011; OLIVEIRA et al., 2012).

No cultivo do cafeeiro, no qual tradicionalmente é obedecido o plantio em nível, ou plantios circulares irrigados sob pivô central, tem-se observado, mesmo

que empiricamente, considerável variação da produtividade entre as diferentes faces da linha de plantio (BICALHO et al., 2005; CHAVES et al., 2012; CUSTÓDIO et al., 2013; OLIVEIRA et al., 2012; VAAST et al., 2006).

De acordo com Bicalho et al. (2005) a face da planta que recebe sol pela manhã tende a apresentar uma maior produção, chegando ao ponto de as duas faces da planta se comportarem até mesmo como lavouras diferentes. Nesse sentido, Oliveira et al. (2012) afirmam que o alinhamento ideal é aquele em que obtém-se uma uniformidade na distribuição da radiação solar em ambas as faces da linha de plantio.

Os frutos do cafeeiro passam por diferentes estágios ao longo de seu desenvolvimento e, até atingirem seu formato final, identifica-se um padrão de crescimento duplo sigmoidal (BELAN et al., 2011; CASTRO; MARRACCINI, 2006; GEROMEL et al., 2006; LAVIOLA et al., 2007; MORAIS et al., 2008; RENA; MAESTRI, 1985).

Para descrever este crescimento duplo sigmoidal será utilizado o modelo não linear duplo Logístico. Esta classe de modelos tem sido muito utilizada na literatura, pois consegue resumir as informações de um conjunto de dados em poucos parâmetros com interpretação prática direta (CARNEIRO et al., 2014; PEREIRA et al., 2014). Um destes parâmetros representa o peso à maturidade do fruto e, como comentado por Bebbington et al. (2009), esta é uma informação de extrema importância para os produtores, principalmente no caso do café por ser comercializado de acordo com seu peso.

Existem diferentes maneiras para se estimar os parâmetros de um modelo de regressão. A metodologia bayesiana deve ser preferida por apresentar melhores resultados e ainda possibilitar a comparação direta entre parâmetros de curvas de crescimento provenientes de diferentes tratamentos (SILVA et al., 2005). A ideia

básica da inferência bayesiana é combinar o conhecimento já existente sobre o fenômeno em estudo (à priori) com o conhecimento presente nos dados (verossimilhança), a fim de obter uma atualização (à posteriori) deste conhecimento por meio do teorema de Bayes.

Devido às expressões dos modelos serem complexas, em geral, a distribuição marginal dos parâmetros não possui uma forma conhecida tornando-se necessário a utilização de técnicas de simulação computacional para gerar amostras destas distribuições e, assim, fazer inferências sobre cada parâmetro. A reamostragem por importância ponderada é uma boa ferramenta para obter amostras das distribuições a posteriori marginais, pois além de ser de fácil compreensão, exige menos esforço computacional do que os algoritmos de Monte Carlo via Cadeias de Markov - MCMC (COLE et al., 2012; SMITH; GELFAND, 1992).

Diante do exposto, o objetivo deste trabalho foi modelar a curva de crescimento dos frutos do cafeeiro, em três alinhamentos de plantio e nas diferentes faces de incidência solar da planta (Norte e Sul), comparando o efeito destes fatores nas estimativas dos parâmetros da curva.

3.2 MATERIAL E MÉTODOS

Os dados foram extraídos de Cunha e Volpe (2011) e correspondem aos resultados de um experimento conduzido na Faculdade de Ciências Agrárias e Veterinárias, UNESP, Campus Jaboticabal, SP (altitude: 615m), com a cultivar Obatã IAC 1669-20, que pertence a espécie *C. arabica*. O experimento foi realizado na safra 2007/2008, na qual as plantas do cafeeiro se encontravam com 5 anos de idade, altura média de 2m e espaçamento de 3,5x0,5m.

Foram analisados três alinhamentos de plantio: I - (81° - 261°), II - (51° - 231°) e III - (66° - 246°) e as coletas feitas separadamente nas faces norte (N) e sul (S) da planta. Cada alinhamento corresponde a um tratamento com

4 repetições, num total de 12 parcelas, com 100 plantas em cada uma, segundo esquema apresentado na Figura 3.1.

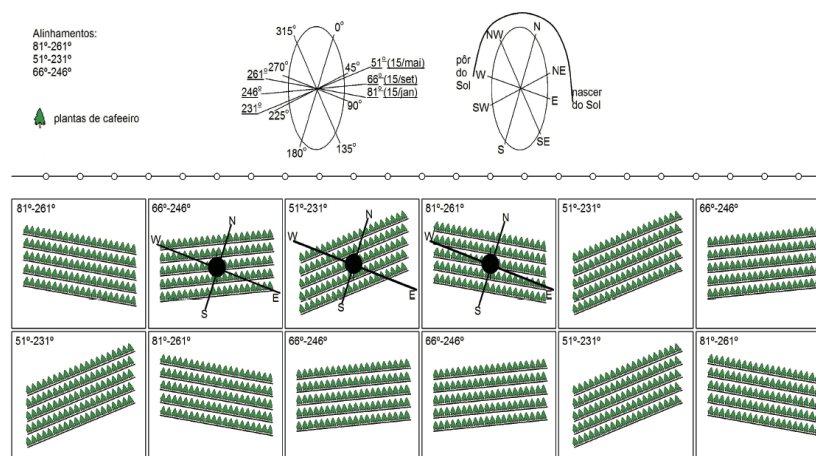


Figura 3.1 Representação esquemática da distribuição dos alinhamentos de plantio e das parcelas do experimento (CUNHA; VOLPE, 2011).

As coletas foram feitas aleatoriamente nas parcelas de um mesmo alinhamento, sendo colhidos 100 frutos de cada alinhamento (50 da face norte e 50 da face sul), totalizando 300 frutos em toda a área do ensaio para cada coleta. Acompanhou-se o desenvolvimento do fruto realizando 14 coletas com um espaçamento médio entre elas de 15 dias, de modo que nestas datas os frutos apresentavam as seguintes idades: 96, 111, 126, 141, 158, 174, 189, 204, 219, 235, 250, 265, 279 e 293 dias após o florescimento - DAF. Os frutos eram acondicionados em sacos plásticos para posterior pesagem em balança analítica de 0,0001g de precisão e assim obter a massa fresca média (em gramas).

As expressões de modelos que descrevem um crescimento duplo sigmoide podem ser obtidas pela soma de dois sigmóides simples, sendo cada um responsável por explicar determinada etapa do crescimento. Nesse sentido o modelo duplo Logístico é obtido pela soma de duas expressões simples do modelo

Logístico. Segundo Fernandes et al. (2015) a parametrização mais adequada é a apresentada abaixo, pois todos os parâmetros apresentam interpretação prática:

$$y_i = \frac{a_1}{1 + e^{k_1(b_1 - x_i)}} + \frac{a_2 - a_1}{1 + e^{k_2(b_2 - x_i)}} + \varepsilon_i \quad (3.1)$$

em que: i representa as observações longitudinais $i = 1, 2, \dots, 14$; y_i é a i -ésima observação da massa fresca média dos frutos; x_i são as idades em dias após o florescimento (DAF) nas quais foram mensuradas a massa fresca média; a_1 representa a massa fresca máxima esperada para a primeira etapa de crescimento; a_2 é a massa fresca final, ou peso à maturidade do fruto; b_1 e b_2 representam a abscissa do ponto de inflexão nas respectivas etapas; k_1 e k_2 representam o índice de crescimento nas etapas 1 e 2, quanto maior valor de k mais precoce é o fruto e menos tempo é necessário para que ele atinja seu valor máximo (a_1 ou a_2); e ε_i é o resíduo de ajuste do modelo o qual pressupõe-se que seja distribuído da seguinte forma $\varepsilon \sim N(0, \sigma^2)$.

Portanto o vetor de parâmetros é dado por $\theta' = [a_1, b_1, k_1, a_2, b_2, k_2]$. Como já comentado, todos os parâmetros apresentam interpretação prática e o mais importante é o peso à maturidade do fruto, a_2 (BEBBINGTON et al., 2009). Posteriormente, após a obtenção das estimativas dos parâmetros para cada face e cada alinhamento foi realizada uma comparação entre os valores estimados destes parâmetros para verificar se existe diferença na curva de crescimento entre as faces dos alinhamentos de plantio e entre os próprios alinhamentos.

Com a interpretação prática dos parâmetros, indicar distribuições à priori para cada um deles não foi muito difícil, pois conhecendo o significado dos parâmetros, é possível estabelecer um intervalo de valores plausíveis para cada um. Desta forma, utilizou-se a distribuição beta como priori para todos os parâmetros, uma vez que esta é indicada para modelar variáveis aleatórias cujas realizações

estejam restritas a um intervalo identificável. Assim as prioris de todos os parâmetros possuem distribuição de probabilidade beta ($P(\theta_i) \sim Beta(\alpha_i, \beta_i)$) em que α_i e β_i são os hiperparâmetros.

Com base no conhecimento sobre o modelo duplo Logístico e principalmente sobre o crescimento do fruto do cafeeiro foram definidas as seguintes prioris: [Beta(2,2)], [200*Beta(2,2)], [Beta(2,10)], [2*Beta(2,2)], [200*Beta(2,2) + 200] e [Beta(2,20)] para os parâmetros a_1, b_1, k_1, a_2, b_2 e k_2 respectivamente. Assumindo independência entre as prioris, a distribuição à priori conjunta para o vetor de parâmetros fica:

$$P(\theta) \propto [a_1(1-a_1)][200b_1(1-b_1)][k_1(1-k_1)^9][2a_2(1-a_2)][200b_2(1-b_2)+200][k_2(1-k_2)^{19}].$$

Sob a pressuposição de normalidade residual, a expressão da verossimilhança do modelo é dada por:

$$L(Y|\theta) = \prod_{i=1}^n \sqrt{\frac{1}{2\pi\sigma^2}} \exp \left\{ \frac{-1}{2\sigma^2} \left[y_i - \frac{a_1}{1 + e^{k_1(b_1-x_i)}} - \frac{a_2 - a_1}{1 + e^{k_2(b_2-x_i)}} \right]^2 \right\}.$$

Representando a expressão da priori por $P(\theta)$ e fazendo a combinação dela com a verossimilhança do modelo tem-se a distribuição à posteriori:

$$P(\theta|Y) \propto \exp \left\{ \frac{-1}{2\sigma^2} \sum_{i=1}^n \left[y_i - \frac{a_1}{1 + e^{k_1(b_1-x_i)}} - \frac{a_2 - a_1}{1 + e^{k_2(b_2-x_i)}} \right]^2 \right\} \times P(\theta).$$

As informações sobre as distribuições marginais de cada parâmetro foram obtidas via reamostragem por importância ponderada. Foi escolhida como distribuição candidata uma aproximação da distribuição de verossimilhança, o que simplifica o cálculo dos pesos uma vez que estes podem ser obtidos diretamente da distribuição à priori. Foram obtidos também os intervalos de credibilidade de máxima densidade a posteriori (do inglês HPD) com o nível de credibilidade de

95%.

De acordo com Silva et al. (2005), em um contexto bayesiano, a comparação entre parâmetros de curvas diferentes pode ser feita de maneira simples e direta. Os autores sugerem que sejam geradas amostras de uma distribuição marginal para a diferença entre os parâmetros de interesse. Posteriormente calcula-se o intervalo de máxima densidade a posteriori (HPD) para esta diferença, caso o intervalo gerado contenha o valor zero, os parâmetros são considerados iguais, caso contrário, são considerados diferentes.

Deste modo, com o intuito de comparar a curva de crescimento dos frutos entre as faces de exposição ao sol (norte, sul), foram geradas cadeias para a diferença entre os parâmetros das faces dentro de cada alinhamento. Para comparar o efeito de alinhamento, foi ajustado um modelo para cada um dos três alinhamentos, desconsiderando o efeito das faces, e geradas cadeias para a diferença entre os novos parâmetros.

O processo de reamostragem, bem como toda a parte computacional envolvida na elaboração dos resultados deste trabalho foram feitos utilizando o software estatístico R (R CORE TEAM, 2015). Os gráficos e as medidas descritivas, como média e desvio padrão das cadeias, foram obtidos utilizando o pacote *coda* também deste software.

3.3 RESULTADOS E DISCUSSÃO

A distribuição conjunta à posteriori ($P(\theta|Y)$) possui uma expressão complexa, tornando-se impossível obter um formato conhecido para as distribuições marginais de cada parâmetro ($\theta' = [a_1, b_1, k_1, a_2, b_2, k_2]$). Utilizou-se o método de reamostragem por importância ponderada, gerando amostras de tamanho $n = 10000$ para cada distribuição marginal. São apresentadas nas Tabelas 3.1, 3.2 e 3.3 as médias das distribuições marginais a posteriori, o desvio padrão e os intervalos

HPD (95%) para os parâmetros do modelo duplo Logístico na descrição da curva de crescimento dos frutos do cafeeiro em ambas as faces dos alinhamentos de plantio I, II e III, respectivamente.

Tabela 3.1 Valores médios com os respectivos desvios padrão e limites inferior (LI) e superior (LS) do intervalo de credibilidade para os parâmetros do modelo duplo Logístico na descrição da curva de crescimento de frutos do cafeeiro no alinhamento de plantio 81° 261°.

Parâmetro	média	Desvio padrão	HPD 95%	
			LI	LS
NORTE				
a_1	0,6708	0,0418	0,5863	0,7494
b_1	148,126	1,5751	145,06	151,192
k_1	0,1659	0,0281	0,1099	0,2194
a_2	1,0817	0,0264	1,0316	1,1336
b_2	225,601	5,1765	215,151	235,788
k_2	0,0631	0,0196	0,0239	0,0998
SUL				
a_1	0,6432	0,0497	0,5471	0,7401
b_1	146,832	1,8908	142,982	150,42
k_1	0,1731	0,0405	0,0931	0,2512
a_2	1,0361	0,0266	0,9877	1,0906
b_2	220,506	6,2337	208,716	233,1
k_2	0,0651	0,0223	0,0221	0,1083

Embora a diferença não seja estatisticamente significativa, pois os intervalos HPD (95%) possuem interseções, percebe-se que no alinhamento de plantio 81° 261° a face sul tende a ser mais precoce que a face norte, com a curva de crescimento atingindo os pontos de inflexão antes que a face norte. Em outras palavras, as estimativas da média para os parâmetros b_1 e b_2 são menores na face sul, quando comparados com as respectivas da face norte, conforme Tabela 3.1.

No alinhamento de plantio 51° 231° (Tabela 3.2), a estimativa da idade em que a curva de crescimento dos frutos atinge o ponto de inflexão é menor

para a face norte, ao contrário do que ocorre no alinhamento de plantio $81^\circ 261^\circ$. Corroborando assim os resultados de Bicalho et al. (2005) no qual os autores afirmam que as duas faces de uma planta podem se comportar como lavouras diferentes devido a diferença de luminosidade recebida por cada uma delas.

Tabela 3.2 Valores médios com os respectivos desvios padrão e limites inferior (LI) e superior (LS) do intervalo de credibilidade para os parâmetros do modelo duplo Logístico na descrição da curva de crescimento de frutos do cafeeiro no alinhamento de plantio $51^\circ 231^\circ$.

Parâmetro	média	Desvio padrão	HPD 95%	
			LI	LS
NORTE				
a_1	0,6077	0,0722	0,4711	0,7487
b_1	146,288	2,3817	141,753	151,045
k_1	0,1443	0,0385	0,0689	0,2198
a_2	1,0619	0,0419	0,9769	1,1414
b_2	224,925	8,1056	208,582	240,564
k_2	0,0537	0,0201	0,0139	0,0916
SUL				
a_1	0,6499	0,0390	0,5775	0,7315
b_1	150,476	2,0061	146,315	154,208
k_1	0,0997	0,0137	0,0743	0,1266
a_2	1,0536	0,0306	0,9961	1,1164
b_2	238,167	4,8871	228,613	247,713
k_2	0,0624	0,0202	0,0221	0,1001

Conforme comentado por Bebbington et al. (2009) o peso à maturidade dos frutos possui grande importância para os produtores, portanto foi dada atenção especial ao parâmetro a_2 , o qual estima tal informação. Este parâmetro foi muito bem estimado, pois o erro padrão foi pequeno e ainda todos os intervalos de credibilidade contém as últimas medições (que conseqüentemente representam o peso final dos frutos) de massa fresca observadas por Cunha e Volpe (2011).

Tabela 3.3 Valores médios com os respectivos desvios padrão e limites inferior (LI) e superior (LS) do intervalo de credibilidade para os parâmetros do modelo duplo Logístico na descrição da curva de crescimento de frutos do cafeeiro no alinhamento de plantio $66^\circ 246^\circ$.

Parâmetro	média	Desvio padrão	HPD 95%	
			LI	LS
NORTE				
a_1	0,4759	0,1204	0,2462	0,7115
b_1	149,53	3,3432	143,147	156,078
k_1	0,1329	0,0539	0,0375	0,2439
a_2	1,0443	0,0404	0,9639	1,1225
b_2	208,764	11,9111	184,635	230,963
k_2	0,0403	0,0089	0,0221	0,0567
SUL				
a_1	0,4986	0,0580	0,3796	0,6064
b_1	147,911	2,0128	144,11	151,957
k_1	0,2017	0,0619	0,0805	0,3212
a_2	1,0547	0,0405	0,9754	1,1336
b_2	216,201	9,8106	196,875	235,408
k_2	0,0317	0,0049	0,0222	0,0415

O desvio padrão das distribuições marginais para os parâmetros do alinhamento de plantio $81^\circ 261^\circ$ (Tabela 3.1), em geral, foi menor quando comparado com os das respectivas distribuições nos demais alinhamentos (Tabelas 3.2 e 3.3), sugerindo assim cadeias mais homogêneas e menor variabilidade entre os frutos. Observando a Figura 3.1 percebe-se que este alinhamento é o menos perpendicular à trajetória aparente do sol, ou seja, neste a luz solar fica mais bem dividida entre as faces durante o dia corroborando assim com Oliveira et al. (2012) que afirmaram que o alinhamento ideal é aquele que obtém uniformidade na radiação solar em ambas as faces.

Objetivando comparar o peso a maturidade do fruto entre as faces de um mesmo alinhamento foram geradas amostras das distribuições Δ_1 , Δ_2 e Δ_3 , as

quais correspondem às diferenças marginais entre os parâmetros a_2 das faces norte e sul nos três alinhamentos, respectivamente. Depois, conforme sugerido por Silva et al. (2005), calculou-se o intervalo de máxima densidade a posteriori de 95% de credibilidade (HPD - 95%), cujos resultados estão apresentados na Tabela 3.4.

Ainda na Tabela 3.4 estão apresentados também o resultado da comparação do peso a maturidade dos frutos entre os alinhamentos, sendo que as distribuições Δ_4 , Δ_5 e Δ_6 correspondem às diferenças marginais entre os parâmetros dos a_2 dos três alinhamentos comparados, obtidas de maneira semelhante a comparação entre faces.

Tabela 3.4 Resumo das distribuições da diferença marginal do peso à maturidade estimado (a_2) entre as faces de um mesmo alinhamento (Δ_1 , Δ_2 e Δ_3) e entre os alinhamentos (Δ_4 , Δ_5 e Δ_6).

Alinhamento	Parâmetro	média	Desvio padrão	HPD 95%	
				LI	LS
I = 81° 261°	Δ_1	0,0452	0,0368	-0,0290	0,1157
II = 51° 231°	Δ_2	0,0075	0,0515	-0,0894	0,1114
III = 66° 246°	Δ_3	-0,0108	0,0567	-0,1201	0,1022
I com II	Δ_4	0,0009	0,0437	-0,0797	0,0088
I com III	Δ_5	0,0141	0,0465	-0,0735	0,1075
II com III	Δ_6	0,0129	0,0511	-0,0905	0,1085

Nenhuma das diferenças, tanto entre faces quanto entre os alinhamentos, foi significativa ao nível de 95% de credibilidade, isto é, todos os intervalos HPD (95%) para as diferenças contém o valor zero. Portanto, não existem evidências estatísticas para afirmar que o peso à maturidade do fruto seja diferente entre as curvas de crescimento estudadas. Embora estão apresentadas na Tabela 3.4 apenas os resultados para o peso à maturidade dos frutos (parâmetro mais importante), foram comparados da mesma forma todos os demais parâmetros do modelo duplo Logístico e todos os intervalos de credibilidade obtidos contém o valor zero. Contradizendo assim a hipótese inicial de que as orientações de plantio, bem como

as faces de exposição ao sol interferem na curva de crescimento do fruto.

Estudando a curva de crescimento do fruto do cafeeiro em diferentes alinhamentos de plantio, embora considerando um modelo sigmoidal simples e sem fazer a comparação direta entre os parâmetros, Fernandes et al. (2014) também indicaram um mesmo modelo de crescimento para ambas as faces, sugerindo portanto não haver diferença entre o padrão de crescimento, assim como observado neste trabalho.

Autores como Bicalho et al. (2005), Chaves et al. (2012), Custodio et al. (2013), Oliveira et al. (2012) e Vaast et al. (2006) encontraram diferença na produtividade entre faces da planta. No entanto, esta diferença em produtividade não implica em diferença no padrão de crescimento dos frutos, assim os resultados aqui obtidos não contradizem necessariamente o que foi observado por estes autores. Principalmente porque, conforme Cunha e Volpe (2011), foi coletada a mesma quantidade de frutos, aleatoriamente, em ambas as faces da planta, não considerando se em determinada face havia uma maior produtividade.

Custódio et al. (2015) observaram além da diferença de produtividade uma variação também no aspecto sensorial do café. Avaliando a qualidade da bebida Barbosa et al. (2010), Castro e Marraccini (2006) e Chaves et al. (2012) também encontraram diferença nas componentes químicas entre os frutos de ambas as faces. Tais aspectos não foram avaliados neste trabalho mas não influenciam no padrão de crescimento dos frutos, pois não foram encontradas diferenças entre os parâmetros das curva de crescimento provenientes das diferentes faces nos três alinhamentos de plantio estudados.

Outro aspecto que pode ser observado com base na média das distribuições marginais dos parâmetros é a diferença na intensidade de crescimento entre as duas etapas de desenvolvimento do fruto. Em todas as 6 situações estudadas

as estimativas de k_1 são maiores do que as do respectivo k_2 , sugerindo um desenvolvimento em massa fresca mais acelerado na primeira etapa. Assim, foram geradas também cadeias para a diferença entre k_1 e k_2 nas curvas de crescimento das faces norte (Δ_{kn}) e sul (Δ_{ks}) dos três alinhamentos de plantio estudados e calculou-se os intervalos HPD de 95% de credibilidade.

Tabela 3.5 Resumo das distribuições da diferença marginal dos parâmetros k_1 e k_2 de uma mesma curva de crescimento nas faces norte (Δ_{kn}) e sul (Δ_{ks}) de cada alinhamento.

Alinhamento	Parâmetro	média	Desvio padrão	HPD 95%	
				LI	LS
81° 261°	Δ_{kn}	0,1039	0,0397	0,0301	0,1844
81° 231°	Δ_{ks}	0,1092	0,0537	0,0049	0,2114
51° 231°	Δ_{kn}	0,0925	0,0509	0,0122	0,1792
51° 231°	Δ_{ks}	0,0379	0,0282	0,0057	0,0863
66° 246°	Δ_{kn}	0,0926	0,0576	0,0016	0,1873
66° 246°	Δ_{ks}	0,1694	0,0615	0,0474	0,2763

Os HPDs da Tabela 3.5 não contém o valor zero, comprovando que a diferença entre k_1 e k_2 é estatisticamente significativa com 95% de credibilidade. Além disso, todas as diferenças foram obtidas pelo seguinte cálculo $\Delta = k_1 - k_2$, logo como os HPDs estão todos concentrados em valores positivos pode-se ainda concluir que $k_1 > k_2$. Tal fato pode ser explicado pela tabela de escala fenológica do cafeeiro apresentada por Moraes et al. (2008), na qual o estágio de rápida expansão do fruto, tanto em ganho de massa como em volume, coincide com a primeira etapa do crescimento, ficando a segunda etapa responsável apenas pela retenção de água e mudança da coloração externa que fazem parte do processo de maturação dos frutos.

Desta forma pode-se elucidar uma única curva de crescimento duplo sigmoideal para o fruto do cafeeiro, com todos os dados do experimento e desconsiderando a presença de faces e alinhamentos. As estimativas para os parâmetros

desta curva média estão apresentados na Tabela 3.6 e a curva com a respectiva taxa de crescimento na Figura 3.2. Percebe-se que diferente da ideia concebida por Coombe (1976), na qual o autor sugeriu que a taxa de crescimento seria a mesma nas duas etapas do crescimento, no fruto do cafeeiro, esta taxa é bem maior na primeira etapa. Ao contrário também do que ocorre com o fruto do pêsego, pois neste fruto Silva et al. (2013) observaram um crescimento mais acelerado na segunda etapa.

Tabela 3.6 Valores médios com os respectivos desvios padrão e limites inferior (LI) e superior (LS) do intervalo de credibilidade para os parâmetros do modelo duplo Logístico na descrição da curva de crescimento de frutos do cafeeiro.

Parâmetro	média	Desvio padrão	HPD 95%	
			LI	LS
a1	0,5545	0,0846	0,3491	0,7599
b1	147,3882	1,9506	143,5654	151,2113
k1	0,1614	0,0474	0,0684	0,2544
a2	1,0677	0,0445	0,9803	1,1551
b2	219,4341	8,3522	201,1442	240,7245
k2	0,0411	0,0155	0,0106	0,0715

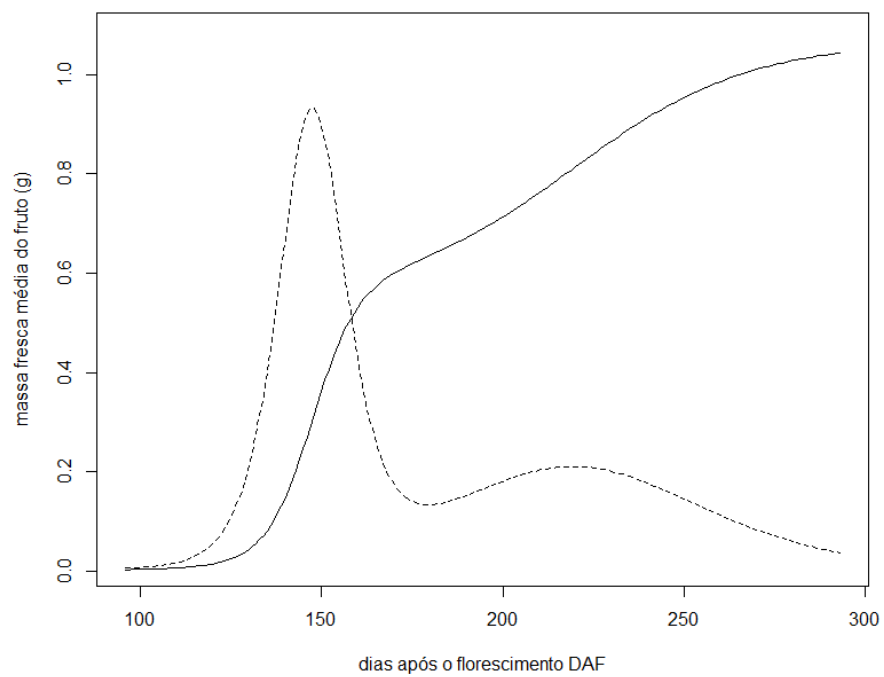


Figura 3.2 Curva de crescimento em acúmulo de massa fresca (—) e taxa de crescimento (- - -) de frutos do cafeeiro.

3.4 CONCLUSÕES

Existe diferença na velocidade de crescimento entre as duas etapas da curva duplo sigmoidal que descreve o padrão de crescimento do fruto do cafeeiro, independente da face e do alinhamento, sendo que em todas as situações a primeira etapa é mais acelerada, com um crescimento mais evidente.

Embora seja consolidado na literatura que existe diferença na qualidade da bebida e principalmente de produtividade entre as faces e alinhamentos de plantio do cafeeiro, a curva de crescimento dos frutos não é influenciada por estes fatores.

REFERÊNCIAS

- BARBOSA, J. N. et al. Spatial distribution of coffees from Minas Gerais state and their relation with quality. **Coffee Science**, Lavras, v. 5, n. 3, p. 237-250, set./dez. 2010.
- BEBBINGTON, M. et al. Dynamics and phases of kiwifruit (*Actinidia deliciosa*) growth curves. **New Zealand Journal of Crop and Horticultural Science**, Wellington, v. 37, n. 3, p. 179-188, July 2009.
- BELAN, L. L. et al. Aspectos fisiológicos do cafeeiro conilon: uma abordagem sistemática. **Nucleus**, Ituverava, v. 8, n. 1, p. 225-240, abr. 2011.
- BICALHO, G. O. D. et al. Direcionamento das linhas de plantio em diferentes orientações cardiais e seus reflexos sobre a produtividade de cafeeiros. In: SIMPÓSIO DE PESQUISA DOS CAFÉS DO BRASIL, 2005, Londrina. **Anais...** Brasília: EMBRAPA Café, 2005. Disponível em: <<http://www.sbicafe.ufv.br/handle/123456789/2107>>. Acesso em: 5 jan. 2016.
- CARNEIRO, A. P. S. et al. Identidade de modelos não lineares para comparar curvas de crescimento de bovinos da raça Tabapuã. **Pesquisa Agropecuária Brasileira**, Brasília, v. 49, n. 1, p. 57-62, jan. 2014.
- CASTRO, R. D.; MARRACCINI, P. Cytology, biochemistry and molecular changes during coffee fruit development. **Brazilian Journal of Plant Physiology**, Campos dos Goytacazes, v. 18, n. 1, p. 175-199, Jan./Mar. 2006.
- CHAVES, A. R. M. et al. Varying leaf-to-fruit ratios affect branch growth and dieback, with little to no effect on photosynthesis, carbohydrate or mineral pools, in different canopy positions of field-grown coffee trees. **Environmental and Experimental Botany**, Elmsford, v. 77, p. 207-218, Apr. 2012.
- COLE, S. R. et al. Bayesian posterior distributions without markov chains. **American Journal of Epidemiology**, Oxford, v. 175, n. 5, p. 368-375, Feb. 2012.
- COOMBE, B. G. The development of fleshy fruits. **Annual Review of Plant Physiology**, Palo Alto, v. 27, p. 507-528, 1976.

- CUNHA, A. R.; VOLPE, C. A. Curvas de crescimento do fruto de cafeeiro cv. Obatã IAC 1669-20 em diferentes alinhamentos de plantio. **Semina: Ciências Agrárias**, Londrina, v. 32, n. 1, p. 49-62, jan./mar. 2011.
- CUSTODIO, A. A. P. et al. Características vegetativas e produtivas de cafeeiros em diferentes supressões de irrigação e faces de exposição. **Coffee Science**, Lavras, v. 8, n. 4, p. 411-422, out./dez. 2013.
- CUSTODIO, A. A. P. et al. Qualidade do café sob manejos de irrigação, faces de exposição solar e posições na planta. **Irriga**, Botucatu, v. 20, n. 1, p. 177-192, jan./mar. 2015.
- FERNANDES, T. J. et al. Parameterization effects in nonlinear models to describe growth curves. **Acta Scientiarum. Technology**, Maringá, v. 37, n. 4, p. 397-402, Oct./Dec. 2015.
- FERNANDES, T. J. et al. Seleção de modelos não lineares para descrição das curvas de crescimento do fruto do cafeeiro. **Coffee Science**, Lavras, v. 9, n. 2, p. 207-215, abr./jun. 2014.
- GEROMEL, C. et al. Biochemical and genomic analysis of sucrose metabolism during coffee (*Coffea arabica*) fruit development. **Journal of Experimental Botany**, Lancaster, v. 57, n. 12, p. 3243-3258, Aug. 2006.
- LAVIOLA, B. G. et al. Acúmulo de nutrientes em frutos de cafeeiro em quatro altitudes de cultivo: cálcio, magnésio e enxofre. **Revista Brasileira de Ciência do Solo**, Campinas, v. 31, p. 1451-1462, maio 2007.
- MORAIS, H. et al. Escala fenológica detalhada da fase reprodutiva de *Coffea arabica*. **Bragantia**, Campinas, v. 67, n. 1, p. 693-699, abr. 2008.
- OLIVEIRA, K. M. et al. Modelagem para a estimativa da orientação de linhas de plantio de cafeeiros. **Engenharia Agrícola**, Jaboticabal, v. 32, n. 2, p. 293-305, 2012.
- PEREIRA, A. A. et al. Descrição do crescimento vegetativo do cafeeiro cultivar Rubi MG 1192, utilizando modelos de regressão. **Coffee Science**, Lavras, v. 9, n. 2, p. 236-271, abr./jun. 2014.

R DEVELOPMENT CORE TEAM. **R**: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2015. Disponível em: <<http://www.r-project.org>>. Acesso em: 21 ago. 2015.

RENA, A. B.; MAESTRI, M. Fisiologia do cafeeiro. **Informe Agropecuário**, Belo Horizonte, v. 11, p. 26-40, 1985.

SILVA, A. C. et al. Características produtivas do cafeeiro arábica irrigado por pivô central na região de Lavras, MG. **Coffee Science**, Lavras, v. 6, n. 2, p. 128-136, maio/ago. 2011.

SILVA, D. F. P. et al. Curva de crescimento e padrão respiratório de frutos de genótipos de pessegueiro em região de clima subtropical. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 35, n. 2, p. 642-649, jun. 2013.

SILVA, F. F. et al. Abordagem Bayesiana da curva de lactação de cabras Saanen de primeira e segunda ordem de parto. **Pesquisa Agropecuária Brasileira**, Brasília, v. 40, n. 1, p. 27-33, jan. 2005.

SMITH, A. F. M.; GELFAND, A. E. Bayesian statistics without tears: a sampling-resampling perspective. **The American Statistician**, Washington, v. 46, n. 2, p. 84-88, May 1992.

VAAST, P. et al. Fruit thinning and shade improve bean characteristics and beverage quality of coffee (*Coffea arabica* L.) under optimal conditions. **Journal of the Science of Food and Agriculture**, London, v. 86, n. 2, p. 197-204, Jan. 2006.