Universidade de São Paulo
Escola Superior de Agricultura “Luiz de Queiroz”

Modelos não lineares resultantes da soma de regressões lineares
ponderadas por funções distribuição acumulada

Lucas Santana da Cunha

Tese apresentada para obtenção do título de
Doutor em Ciências. Área de concentração:
Estatística e Experimentação Agronômica

Piracicaba
2016

Lucas Santana da Cunha
Licenciado em Matemática

**Modelos não lineares resultantes da soma de regressões lineares
ponderadas por funções distribuição acumulada**

versão revisada de acordo com a resolução CoPGr 6018 de 2011

Orientador:
Prof^a. Dr^a. **SÔNIA MARIA DE STEFANO PIEDADE**

Tese apresentada para obtenção do título de
Doutor em Ciências. Área de concentração: Es-
tatística e Experimentação Agronômica

**Piracicaba
2016**

**Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA - DIBD/ESALQ/USP**

Cunha, Lucas Santana da

Modelos não lineares resultantes da soma de regressões lineares ponderadas por funções distribuição acumulada / Lucas Santana da Cunha. - - versão revisada de acordo com a resolução CoPGr 6018 de 2011. - - Piracicaba, 2016.

85 p. : il.

Tese (Doutorado) - - Escola Superior de Agricultura "Luiz de Queiroz".

1. Regressão não linear 2. Modelos de crescimento difásicos 3. Função de distribuição acumulada 4. Agricultura de precisão 5. Controladores eletrônicos de pulverização I. Título

CDD 519.536
C972m

"Permitida a cópia total ou parcial deste documento, desde que citada a fonte – O autor"

DEDICATÓRIA

*Aos meus pais,
Celso Pereira da Cunha e
Eli Terezinha Santana da Cunha
pelo carinho e bons exemplos que me
propiciaram.*

Com amor, DEDICO.

AGRADECIMENTOS

A Deus por sempre estar presente em cada momento de minha vida, me abençoando e dando forças nas dificuldades para realizar este trabalho.

A minha família, em especial aos meus pais Celso Pereira da Cunha e Eli Terezinha Santana da Cunha, ao meu irmão Tiago Santana da Cunha, aos meus avós Armindo Teodoro de Santana, Enelzira Fernandes de Santana, Osvaldo Bueno da Cunha e Arlete Pereira da Cunha, obrigado pelo carinho e força.

A minha namorada Jéssica Costa Teodoro, pela dedicação, compreensão e força nos momentos finais da realização deste trabalho.

A minha orientadora professora Dra. Sônia Maria De Stefano Piedade, pelo apoio, por sempre acreditar em mim, e pela amizade.

Aos professores doutores Carlos Tadeu dos Santos Dias, Clarice Garcia Borges Demetrio, Cristian Marcelo Villegas Logos, Paulo Justiniano, Renata Alcarde, Roseli Aparecida Leandro, Silvio Sandoval Zocchi e Taciana Villela Savian por compartilhar seus conhecimentos e pela amizade.

Aos secretários Rosni Honofre Aparecido Pinto, Luciane Brajão, Mayara Segatto e Solange de Assis Paes Sabadin. Aos técnicos em informática Eduardo Bonilha e Jorge Alexandre Wiendl pelo Suporte.

Ao professor doutor Casimiro Dias Gadanha Junior e ao amigo Ranieri Rodrigues Vieira pela concessão dos dados para realização deste trabalho.

A todos os amigos da Pós-Graduação do LCE da ESALQ/USP e aos amigos de república pela amizade e parceria sempre.

A CAPES pelo fundamental suporte financeiro concedido.

Muito obrigado a todas as pessoas que contribuíram de alguma forma para a realização deste trabalho.

“Talvez não tenha conseguido fazer o melhor, mas lutei para que o melhor fosse feito. Não sou o que deveria ser, mas Graças a Deus, não sou o que era antes.”

(Marthin Luther King).

SUMÁRIO

RESUMO	11
ABSTRACT	13
LISTA DE FIGURAS	15
LISTA DE TABELAS	17
1 INTRODUÇÃO	18
2 REVISÃO BIBLIOGRÁFICA	20
2.1 Controladores eletrônicos de pulverização	20
2.2 Função distribuição acumulada	20
2.2.1 Função distribuição acumulada Logística	20
2.2.2 Função distribuição acumulada Gumbel	21
2.2.3 Função distribuição acumulada Cauchy seno hiperbólico exponencial	22
2.3 Modelos de regressão linear simples	22
2.4 Modelos de regressão não linear	23
2.4.1 Modelo de regressão não linear logístico	24
2.4.2 Modelo de regressão não linear de Gompertz	25
2.5 Reparametrização dos modelos não lineares	25
2.6 Métodos de estimação dos modelos não lineares	26
2.6.1 Verossimilhança	26
2.6.2 Aproximação quadrática da verossimilhança	26
2.7 Seleção de modelos	27
2.7.1 Medidas de curvatura	27
2.7.2 Critério AIC e BIC	27
3 MODELOS DA SOMA DE REGRESSÕES LINEARES PONDERADAS POR FUNÇÕES DE DISTRIBUIÇÃO ACUMULADA	29
3.1 Modelo geral	29
3.2 Caracterização dos modelos de regressão de dois regimes	29
3.2.1 Reparametrização no ponto de fração (PF_q)	30
4 MATERIAL E MÉTODOS	32
4.1 Tempos de resposta de um sistema eletrônico de pulverização	32
4.2 Dados simulados	35
5 RESULTADOS E DISCUSSÃO	36

5.1 Avaliação dos tempos de resposta	36
5.2 Avaliação dos dados simulados	41
6 CONSIDERAÇÕES FINAIS	45
6.1 Trabalhos futuros	45
REFERÊNCIAS	47
APÊNDICES	51
ANEXOS	57

RESUMO

Modelos não lineares resultantes da soma de regressões lineares ponderadas por funções de distribuição acumulada

Os controladores eletrônicos de pulverização visam minimizar a variação das taxas de insumos aplicadas no campo. Eles fazem parte de um sistema de controle, e permitem a compensação da variação de velocidade de deslocamento do pulverizador durante a operação. Há vários tipos de controladores eletrônicos de pulverização disponíveis no mercado e uma das formas de selecionar qual o mais eficiente nas mesmas condições, ou seja, em um mesmo sistema de controle, é quantificar o tempo de resposta do sistema para cada controlador específico. O objetivo desse trabalho foi estimar os tempos de resposta para mudanças de velocidade de um sistema eletrônico de pulverização via modelos de regressão não lineares, estes, resultantes da soma de regressões lineares ponderadas por funções distribuição acumulada. Os dados foram obtidos no Laboratório de Tecnologia de Aplicação, localizado no Departamento de Engenharia de Biosistemas da Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, no município de Piracicaba, São Paulo, Brasil. Os modelos utilizados foram o logístico e de Gompertz, que resultam de uma soma ponderada de duas regressões lineares constantes com peso dado pela função distribuição acumulada logística e Gumbell, respectivamente. Reparametrizações foram propostas para inclusão do tempo de resposta do sistema de controle nos modelos, com o objetivo de melhorar a interpretação e inferência estatística dos mesmos. Foi proposto também um modelo de regressão não linear difásico que resulta da soma ponderada de regressões lineares constantes com peso dado pela função distribuição acumulada Cauchy seno hiperbólico exponencial. Um estudo de simulação foi feito, utilizando a metodologia de Monte Carlo, para avaliar as estimativas de máxima verossimilhança dos parâmetros do modelo.

Palavras-chave: Agricultura de precisão; Controladores eletrônicos de pulverização; Tempos de resposta; Regressão não linear; Modelos de crescimento difásico

ABSTRACT

Nonlinear Models resulting from the sum of weighted linear regression of cumulative distribution functions

The electronic controllers spray aimed at minimizing the variation of inputs rates applied in the field. They are part of a control system, and allow for compensation for variation spray travel speed during operation. There are several types of electronic spray controllers on the market and one way to select which more efficient under the same conditions, ie in the same system of control, is to quantify the system response time for each specific driver. The objective of this study was to estimate the response times for changes in speed of an electronic spraying system via nonlinear regression models, these resulting from the sum of weighted linear regressions for cumulative distribution functions. Data were obtained on the Application Technology Laboratory, located in the Department of Biosystems Engineering from College of Agriculture “Luiz de Queiroz”, University of Sao Paulo, in Piracicaba, Sao Paulo, Brazil. The models used were the logistic and Gompertz, resulting from a weighted sum of two constant linear regressions with weight given by the cumulative distribution function logistics and Gumbell respectively. Reparametrization been proposed for inclusion in the control system response time models, in order to improve the statistical interpretation and inference of the same. It has also been proposed a non-linear regression model two-phase which is the weighted sum of constant linear regressions weight given by a cumulative distribution function exponential hyperbolic sine Cauchy in which a simulation study was conducted using the methodology of Monte Carlo to evaluating the maximum likelihood estimates of the model parameters.

Keywords: Precision agriculture; Electronic controllers spray; Response times;
Nonlinear regression; Growth models diphasic

LISTA DE FIGURAS

Figura 1 - Formas da função distribuição acumulada Logística para diferentes valores de seus parâmetros	21
Figura 2 - Formas da função distribuição acumulada Gumbel para diferentes valores de seus parâmetros	21
Figura 3 - Formas da função distribuição acumulada Exp-SC para diferentes valores de seus parâmetros	22
Figura 4 - Pressão (kPa) em função do tempo (s), para a ponta de pulverização XR 110 02 (baixa vazão), de cada mudança de velocidade	32
Figura 5 - Pressão (kPa) em função do tempo (s), para a ponta de pulverização XR 110 04 (média vazão), de cada mudança de velocidade	33
Figura 6 - Pressão (kPa) em função do tempo (s), para a ponta de pulverização XR 110 06 (alta vazão), de cada mudança de velocidade	33
Figura 7 - Valores observados e preditos com banda de confiança assintótica (95%) para a pressão (kPa) em função do tempo (s) para cada mudança de velocidade. Linhas verticais identificam a estimativa intervalar assintótica (95%) do tempo de resposta do controlador eletrônico de pulverização para a ponta de baixa vazão XR 110 02	37
Figura 8 - Valores observados e preditos com banda de confiança assintótica (95%) para a pressão (kPa) em função do tempo (s) para cada mudança de vazão. Linhas verticais identificam a estimativa intervalar assintótica (95%) do tempo de resposta do controlador eletrônico de pulverização para a ponta de média vazão XR 110 04	39
Figura 9 - Valores observados e preditos com banda de confiança assintótica (95%) para a pressão (kPa) em função do tempo (s) para cada mudança de velocidade. Linhas verticais identificam a estimativa intervalar assintótica (95%) do tempo de resposta do controlador eletrônico de pulverização para a ponta de alta vazão XR 110 06	40
Figura 10 - Valores observados para os verdadeiros parâmetros $\beta_{20} = 50$; $x_I = 50$; $s = 5$; $\nu = 0,1$ e $\tau = 0,6$ e para as estimativas médias (EMs), em função da variável independente x para os tamanhos amostrais: $n = 18; 45; 135; 270$	42

Figura 11 - Densidades estimadas de 1000 amostras para $n = 18; 45; 135; 270$ dos parâmetros $\beta_{20} = 50; x_I = 50; s = 5; \nu = 0,1$ e $\tau = 0,6$	43
Figura 12 - Gráfico dos resíduos padronizados vs valores ajustados e o envelope simulado para a ponta de pulverização de XR 110 02 (baixa vazão) dos modelos logístico e de Gompertz	53
Figura 13 - Gráfico dos resíduos padronizados vs valores ajustados e o envelope simulado para a ponta de pulverização de XR 110 04 (média vazão) dos modelos logístico e de Gompertz	54
Figura 14 - Gráfico dos resíduos padronizados vs valores ajustados e o envelope simulado para a ponta de pulverização de XR 110 06 (alta vazão) dos modelos logístico e de Gompertz	55

LISTA DE TABELAS

Tabela 1 - Estimativa (Est.), limite inferior (LI) e limite superior (LS) para o intervalo de confiança assintótico (95%) para o parâmetro de interesse tempo de resposta, t_r , dos modelos Logístico e de Gompertz para a ponta de baixa vazão	36
Tabela 2 - Medidas do critério de informação de Akaike (AIC) e de curvatura devido ao efeito de parametrização (CP) e intrínseca (CI) dos modelos logístico e de Gompertz para a ponta de baixa vazão	36
Tabela 3 - Estimativa (Est.), limite inferior (LI) e limite superior (LS) para o intervalo de confiança assintótico (95%) para o parâmetro de interesse tempo de resposta, t_r , dos modelos selecionados para a ponta de baixa vazão	37
Tabela 4 - Estimativa (Est.), limite inferior (LI) e limite superior (LS) para o intervalo de confiança assintótico (95%) para o parâmetro de interesse tempo de resposta, t_r , dos modelos Logístico e de Gompertz para a ponta de média vazão	37
Tabela 5 - Medidas do critério de informação de Akaike (AIC) e de curvatura devido ao efeito de parametrização (CP) e intrínseca (CI) dos modelos logístico e de Gompertz para a ponta de média vazão	38
Tabela 6 - Estimativa (Est.), limite inferior (LI) e limite superior (LS) para o intervalo de confiança assintótico (95%) para o parâmetro de interesse tempo de resposta, t_r , dos modelos selecionados para a ponta de média vazão	39
Tabela 7 - Estimativa (Est.), limite inferior (LI) e limite superior (LS) para o intervalo de confiança assintótico (95%) para o parâmetro de interesse tempo de resposta, t_r , dos modelos Logístico e de Gompertz para a ponta de alta vazão	40
Tabela 8 - Medidas do critério de informação de Akaike (AIC) e de curvatura devido ao efeito de parametrização (CP) e intrínseca (CI) dos modelos logístico e de Gompertz para a ponta de alta vazão	40
Tabela 9 - Estimativa (Est.), limite inferior (LI) e limite superior (LS) para o intervalo de confiança assintótico (95%) para o parâmetro de interesse tempo de resposta, t_r , dos modelos selecionados para a ponta de alta vazão	41

Tabela 10 -Estimativa (Est.), limite inferior (LI), limite superior (LS) para o intervalo de confiança assintótico (95%) e estimativa de Vieira (2013) (Est. Antes), para o parâmetro de interesse tempo de resposta, t_r , para as três diferentes pontas	41
Tabela 11 -EMs, Vies, EQMs, CA e K baseados em 1000 simulações do modelo de regressão não linear Exp-SC de parâmetros $\beta_{20} = 50$; $x_I = 50$; $s = 5$; $\nu = 0,1$ e $\tau = 0,6$	42
Tabela 12 -Valores da pressão (Kpa) em função do tempo (s) da ponta de pulverização XR 110 02 (baixa vazão) para as mudanças realizadas	56

1 INTRODUÇÃO

O crescimento das áreas de plantio e, principalmente, o aumento da produtividade têm sido associados ao uso de novas tecnologias. Um dos fatores que tem contribuído para o aumento da produtividade na agricultura é o emprego de produtos fitossanitários, os quais têm se mostrado fundamentais para a proteção e preservação do potencial produtivo de culturas agrícolas. O setor agrícola brasileiro é um grande mercado para a indústria mundial de defensivos agrícolas. Em 2014, comercializou o valor total de US\$12,25 bilhões em defensivos, em que a soja foi a principal cultura consumidora de defensivos no país, sendo responsável por 55,6% (US\$6,80 bilhões) do valor total das vendas.

No entanto, tecnologia consiste na aplicação dos conhecimentos científicos a um determinado processo produtivo. Dessa forma, entende-se como “tecnologia de aplicação de produtos fitossanitários” o emprego de todos os conhecimentos científicos que proporcionem a correta colocação do produto biologicamente ativo no alvo, em quantidade necessária, de forma econômica, com o mínimo de contaminação de outras áreas (MATUO et. al., 2002).

Os pulverizadores convencionais são calibrados, previamente, para realizar aplicações com pressão ou vazão, e velocidade de deslocamento constantes na operação, resultando assim, numa taxa de aplicação uniforme. No entanto, sabe-se que é quase impossível manter as condições operacionais pré-estabelecida no campo, devido às mais diversas causas, como por exemplo, condições do terreno, obstáculos. Consequentemente, a distribuição do produto fitossanitário não será feita de forma uniforme no alvo, ocasionando em taxas de aplicação diferentes na extensão da área.

Os controladores eletrônicos de pulverização surgem na tentativa de minimizar essa variação das taxas de aplicação devido às mudanças de velocidade. Eles fazem parte de um sistema de controle, e permitem a compensação da variação de velocidade de deslocamento do pulverizador durante a operação, além de oferecer também controle automático do fechamento de seções e aplicações a taxas variáveis (MOHAMMADZAMANI et. al., 2009).

O tempo de resposta do controlador é definido como o período entre o comando e sua efetiva mudança de vazão nas pontas de pulverização no momento de

aplicação (ANTUNIASSI, 1999). Já o tempo determinado pelo instante da mudança até a estabilização por completa do pulverizador, denomina-se tempo de resposta do sistema de controle.

Há vários tipos de controladores eletrônicos de pulverização disponíveis no mercado. Assim, uma das formas de selecionar qual o mais eficiente nas mesmas condições, ou seja, em um mesmo sistema de controle, é quantificar o tempo de resposta do sistema de controle para cada controlador específico. Logo, encontrar uma relação entre a variável vazão e o tempo, considerando o instante de determinada mudança, seria o ideal para quantificar esse tempo de resposta do sistema de controle.

Vários trabalhos foram realizados para obtenção de estimativas pontuais do tempo de resposta de vários sistemas de controle eletrônico de pulverização (CARVALHO e ANTUNIASSI, 2012; GADANHA JUNIOR, 2000). Porém, estes utilizaram métodos empíricos e descritivos. A utilização destes não é o mais adequado uma vez que não se utiliza um modelo probabilístico para o problema e conseqüentemente, não obtém-se estimativas intervalares para os tempos de resposta.

A técnica estatística que pode-se atingir esse objetivo é a análise de regressão. Basicamente, os modelos de regressão são divididos em duas classes: os lineares e não lineares.

Em geral, a regressão linear é usada para especificação de modelos puramente empíricos. Por exemplo, se a resposta de interesse, usualmente representada por Y , depender de uma única variável independente, X , a partir da representação gráfica de X versus Y , pode-se sugerir possíveis modelos. Na presença de várias variáveis independentes, uma alternativa para o ajuste de um possível modelo é partir, inicialmente, de um modelo completo e avaliar a qualidade do ajuste por meio de diagnósticos de regressão (BELSLEY et al., 1980; COOK e WEISBERG, 1982).

Já os modelos de regressão não lineares são considerados quando existe algum conhecimento prévio sobre o estudo para sustentar que a relação entre resposta e preditor segue uma particular forma funcional. Tal conhecimento pode ser desde uma equação diferencial que remete a um particular modelo, como é o caso de modelos de crescimento, ou simplesmente uma restrição sobre a função, como o de ser monótona, típico de curvas de acúmulo, para a qual pode-se ter várias funções disponíveis. Dentre as vantagens da utilização do modelo de regressão não linear, é que frequentemente existe in-

interpretação para a maioria de seus parâmetros (SCHABENBERGUER e PIERCE, 2002).

Fulton et. al. (2005) e Vieira (2013) utilizaram o modelo de regressão não linear logístico de quatro parâmetros para encontrar o tempo de resposta de um sistema de controle eletrônico de pulverização, mas ambos forneceram apenas estimativas pontuais dos mesmos, obtidas por meio de funções dos parâmetros dos modelos.

Tendo em vista que não há trabalhos em que o parâmetro de interesse, tempo de resposta, está representado nos modelos, e, conseqüentemente, não há inferência estatística aos mesmos, como por exemplo, estimativas intervalares, foi proposto neste trabalho uma metodologia que considere tais modelos. Para isso, utilizou-se parte dos dados da dissertação de mestrado de Vieira (2013), de um sistema de controle composto de controlador eletrônico Vcom 8.4® (VERION AGRICULTURA, 2010). O ferramental estatístico utilizado foi os modelos de regressão não lineares obtidos por um processo que consiste de soma de regressões lineares ponderadas por funções de distribuição acumulada.

Uma outra parte deste trabalho foi propor um modelo de regressão não linear difásico resultante da soma de regressões lineares constantes ponderadas por uma função distribuição acumulada Cauchy seno hiperbólico exponencial e mostrar a consistência e validade do mesmo aplicando-o em dados simulados.

Logo, os objetivos desse trabalho foram: (i) mostrar que a soma ponderada de regressões lineares com pesos dado por funções distribuição acumulada resultam em modelos de regressão não lineares e que dois destes são os modelos logístico e de Gompertz; (ii) reparametrizar esses modelos para inclusão do tempo de resposta de um sistema eletrônico de pulverização como um dos parâmetros e assim, obter estimativas pontuais e intervalares para os mesmos; (iii) propor um modelo de crescimento difásico resultante da ponderação de regressões lineares.

2 REVISÃO BIBLIOGRÁFICA

2.1 Controladores eletrônicos de pulverização

O setor agrícola brasileiro é um grande mercado para a indústria mundial de defensivos agrícolas. Em 2014, comercializou o valor total de US\$12,25 bilhões em defensivos, em que a soja foi a principal cultura consumidora de defensivos no país, sendo responsável por 55,6% (US\$6,80 bilhões) do valor total das vendas.

As aplicações dos defensivos realizadas por pulverizadores agrícolas convencionais são operadas a partir do princípio de que a pressão do circuito hidráulico e a velocidade de deslocamento sejam constantes durante a operação, a qual resultará em uma taxa de aplicação uniforme. No entanto, durante as operações de aplicação de produtos fitossanitários a campo, a velocidade de deslocamento não se mantém constante. Ou seja, quando há redução ou aumento de velocidade durante a operação de aplicação, pode ocorrer o aumento ou a diminuição das taxas de aplicação devido ao fato da calibração ser realizada previamente com velocidade constante (ALEXANDER et. al., 1964 apud GADANHA JUNIOR, 2000).

Com a evolução tecnológica da indústria, novos equipamentos possibilitam a correta avaliação e monitoramento das condições de operação realizada no campo, visando o controle efetivo das taxas aplicadas (CARVALHO e ANTUNIASSI, 2012). Tais equipamentos são os controladores eletrônicos de pulverização.

Os controladores visam minimizar a variação das taxas de insumos aplicadas no campo. Esses sistemas permitem a compensação das mudanças de velocidade da máquina durante a operação e atualmente, oferecem grande potencial para realização de controle automático do fechamento de seções e aplicações a taxas variáveis (MOHAMMADZAMANI et. al, 2009).

Eles são compostos por microcomputadores embutidos que possuem entradas para sensores, mapas de prescrição de SIG, interface personalizada de comandos do usuário que por meio de *hardwares* e *softwares* calculam a taxa de aplicação necessária usando seus algoritmos internos. O valor calculado é traduzido e enviado aos atuadores que ajustam as taxas de aplicação, por meio de motores hidráulicos, elétricos e pneumáticos (PRAVEEN, 2009).

Os sistemas de controle baseados na pressão proporcionam a vazão requerida

usando a resposta de um sensor de pressão e os sistemas baseados no fluxo usam a resposta de um medidor de fluxo para corrigir a taxa de aplicação (SHARDA et al., 2011).

O tempo de resposta do controlador é definido como o período entre o comando e sua efetiva mudança de vazão nas pontas de pulverização no momento de aplicação (ANTUNIASSI, 1999). Já o tempo determinado pelo instante da mudança até a estabilização do pulverizador, denomina-se tempo de resposta do sistema de controle.

Para comparação de desempenho de controladores eletrônicos, o tempo de resposta é um dos fatores mais importantes de desempenho, principalmente em sistemas de aplicação localizada e a taxa variada de produtos fitossanitários para agricultura de precisão. É um parâmetro importante tanto para os desenvolvedores do controlador quanto aos seus usuários. Imediatamente, a quantificação destes tempos pode permitir a correção pós-processamento dos mapas de aplicação (FULTON et. al., 2003).

Gadanha Junior (2000) avaliou o desempenho de cinco sistemas de controle eletrônicos de pulverização pelos tempos de resposta dos sistemas durante mudanças de velocidade.

Fulton et al. (2005), Vieira (2013), utilizaram o modelo logístico de quatro parâmetros para encontrar o tempo de resposta de vários aplicadores VRT granulares.

2.2 Função distribuição acumulada

A função distribuição acumulada (FDA), ou função distribuição, descreve a relação entre os valores de uma variável aleatória e sua probabilidade. Assim, uma FDA de uma variável aleatória Y , no caso contínuo, é uma função que a cada número real y associa o valor:

$$F(y) = \mathbb{P}(Y \leq y), \quad x \in \mathbb{R} \quad (1)$$

Assim, podemos observar que a FDA tem como domínio os números reais ($\mathbb{R}$) e imagem o intervalo $[0, 1]$.

A FDA de uma variável aleatória Y tem três propriedades básicas:

1. $0 \leq F(y) \leq 1$, $\lim_{y \rightarrow -\infty} F(y) = 0$ e $\lim_{y \rightarrow \infty} F(y) = 1$;
2. F é não decrescente.
3. F é uma função contínua à direita e tem limite à esquerda.

Essas propriedades específicas caracterizam essa classe de funções, que possuem esse nome por estarem diretamente ligadas a uma variável aleatória Y . Mas, pode-se analisá-las como apenas funções matemáticas que possuem essas características específicas, sem necessariamente estarem ligadas ao conceito de probabilidade. A seguir são apresentadas algumas FDAs e suas respectivas características:

2.2.1 Função distribuição acumulada Logística

A distribuição de probabilidade logística é uma distribuição contínua que assemelha-se a distribuição normal em forma, mas tem caudas mais pesadas (maior curtose). A FDA é:

$$F(x; x_I, s) = \frac{1}{1 + \exp\{-(x - x_I)/s\}}, \quad (2)$$

em que x_I e s são os parâmetros de locação e escala, respectivamente, da função não possuindo parâmetro de forma. Assim, se x_I aumenta ou diminui, a função se desloca para a direita ou para a esquerda, respectivamente. Agora, se s aumenta ou diminui, os valores da função se distanciam ou se aproximam, respectivamente, em direção de x_I . Observa-se da Figura 1 que a função é simétrica em relação ao eixo $x = x_I$.

Figura 1 – Formas da função distribuição acumulada Logística para diferentes valores de seus parâmetros

2.2.2 Função distribuição acumulada Gumbel

Na teoria das probabilidades e estatísticas, a distribuição de Gumbel é contínua e utilizada para modelar a distribuição do máximo (ou o mínimo) de um número de amostras de várias distribuições. A FDA de Gumbel é:

$$F(x; x_I, s) = \exp\{-\exp\{-s(x - x_I)\}\}, \quad (3)$$

em que x_I é o parâmetro de locação e s é o parâmetro de escala, não possuindo parâmetro de forma. Observa-se da Figura 2 que a função é assimétrica à direita.

Figura 2 – Formas da função distribuição acumulada Gumbel para diferentes valores de seus parâmetros

2.2.3 Função distribuição acumulada Cauchy seno hiperbólico exponencial

Para fornecer uma assimetria à distribuição Cauchy seno hiperbólico (SC), Cooray (2013) propôs a distribuição Exp-SC usando a classe exponenciada de distribuições (Gupta e Kundu, 2001), que tem a forma funcional da FDA dada por:

$$F(x; x_I, s, \nu, \tau) = \left\{ \frac{1}{2} + \frac{1}{\pi} \arctan \left[\nu \sinh \left(\frac{x - x_I}{s} \right) \right] \right\}^{\tau}, \quad (4)$$

em que x_I e s são os parâmetros de locação e escala, respectivamente; ν é o parâmetro de simetria, que caracteriza a bimodalidade da distribuição e $\tau > 0$ é o parâmetro de forma.

Figura 3 – Formas da função distribuição acumulada Exp-SC para diferentes valores de seus parâmetros

Devido a maior quantidade de parâmetros, observa-se da Figura 3 que a função é bastante flexível dependendo dos valores de seus parâmetros. Quanto menor o valor do parâmetro ν , a bimodalidade é mais evidente e a curva tem o formato de dois “S”. Agora, se o valor do parâmetro ν aumenta, a função passa torna-se unimodal e a curva tem o formato sigmoidal de apenas um “S”.

2.3 Modelos de regressão linear simples

O termo regressão foi utilizado pela primeira vez, por Sir Francis Galton (1889), que propôs a “Lei de Regressão Universal” em seu livro *Natural Inheritance*. Galton estudava a relação existente entre a estatura dos pais e a estatura dos filhos e

notou que filhos de pais altos tendiam a ser altos, mas não tanto quanto seus pais, e que filhos de pais de menor estatura tendiam a ser baixos, mas também não tanto quanto seus pais. Esse fato foi denominado de regressão para a estatura média (mais precisamente, ele cunhou o termo regressão para a mediocridade). Embora o contexto atual tenha muito pouco a ver com esses primórdios, a denominação permaneceu.

Os modelos de regressão são usados para descrever o comportamento de uma variável aleatória (v.a.) Y como uma função de condicionantes $X_k, k = 1, \dots, p$, chamadas de variáveis explicativas, covariáveis ou estímulos. Com tais modelos busca-se encontrar e descrever padrões de homogeneidade dentre a heterogeneidade observada em um conjunto de observações de Y . Tal procedimento permite (i) explicar, ao menos parcialmente o comportamento de Y ; (ii) quantificar a influência das condicionantes em Y ; (iii) selecionar, mediante algum critério, as condicionantes relevantes; (iv) prever o comportamento de Y para estados observados e também não observados das condicionantes X ; (v) avaliar a incerteza associada a esse processo. O estudo de uma série de problemas práticos, nas mais diversas áreas, podem ser formatado sob esta abordagem, o que faz com que os modelos de regressão sejam largamente adotados e, provavelmente, a mais utilizada dentre as metodologias estatísticas.

O objetivo na adoção de modelos de regressão é associar por meio de uma relação funcional a distribuição (de probabilidades, no caso discreto, ou densidade de probabilidades, no caso contínuo) da v.a. Y aos valores observados de X denotado por $[Y|x]$.

Suponha que a observação Y_i associada a covariável x_i , com $i = 1, \dots, n$, possui distribuição Gaussiana, em que Y_i possui variância constante, e esperança dependendo de x_i pela relação $E[Y_i|x_i] = \mu_i = \beta_0 + \beta_1 x_i$. Este modelo pode ser escrito na forma a seguir.

$$\begin{aligned} Y_i|x_i &\sim N(\mu_i, \sigma^2) \\ \mu_i &= g(\boldsymbol{\beta}, x_i) = \beta_0 + \beta_1 x_i. \end{aligned} \tag{5}$$

O modelo (5) chamado de modelo de regressão linear simples, sendo linear por impor que $g(\boldsymbol{\mu}_i)$ seja uma função linear nos coeficientes (β_0, β_1) , que são os elementos do vetor de parâmetros $\boldsymbol{\beta}$, e simples por tal função envolver apenas uma variável explicativa x . O modelo pode ser escrito de forma alternativa, como:

$$Y_i|x_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

$$\epsilon_i \sim N(0, \sigma^2).$$

Representando a relação $Y_i|x_i$ em um plano cartesiano, verifica-se que para $x_i = 0$, β_0 representa o ponto onde a reta corta o eixo dos Y 's (ordenada) e por isso é chamado intercepto. Já β_1 é chamado coeficiente angular da reta, pois, da interpretação geométrica de derivada tem-se

$$\beta_1 = tg(\alpha)$$

sendo α o ângulo que a reta forma como o eixo dos X 's (abscissa). Além disso, tem-se que para um aumento de 1 unidade de X a $E(Y_i|x_i)$ aumenta β_1 unidades.

Observa-se que para um caso particular, $\alpha = 0$, temos que $\beta_1 = 0$, e assim, a equação (5) passa a ser

$$Y_i|x_i \sim N(\mu_i, \sigma^2)$$

$$\mu_i = g(\beta, x_i) = \beta_0. \quad (6)$$

ou seja, o valor esperado de Y dado x é a constante β_0 para qualquer valor de x_i . Denomina-se esse modelo como regressão linear constante.

2.4 Modelos de regressão não linear

Modelos de regressão não linear usualmente são sustentados por alguma informação sobre a relação entre Y e X . Tal informação está vinculada à diferentes graus de conhecimento como: uma análise de um diagrama de dispersão de y vs x ; restrições de forma da função (ser monótona, ser sigmóide); a solução de uma equação diferencial sustentada por algum princípio/teoria; interpretação dos seus parâmetros. Seja qual for o grau de conhecimento, a escolha de um modelo não linear raramente é empírica.

Segundo Bates e Whatts (1988), Draper e Smith (1998) e Ratkowsky (1983), entre outros, a definição para que um modelo de regressão seja designado não linear é de que pelo menos uma das derivadas parciais de uma função η em relação aos parâmetros depende de pelo menos um dos parâmetros do modelo, ou seja, pode-se dizer que um modelo é não linear quando ele é não linear nos parâmetros.

Estatisticamente, um modelo de regressão não linear, descreve alguma quantidade relacionada à distribuição de probabilidades de uma variável aleatória Y como uma função não linear nos parâmetros. Na grande maioria das vezes, a quantidade de interesse é a esperança de Y . Dessa forma,

$$\mu = E(Y|x) = \eta(x, \boldsymbol{\theta})$$

representa um modelo que descreve, como quantidade de interesse, a esperança de uma variável aleatória Y em termos da variável explicativa X e do vetor de p parâmetros $\boldsymbol{\theta} = (\theta_1, \dots, \theta_p)^T$ por meio da função não linear η .

Os modelos de regressão não linear geralmente são utilizados para respostas contínuas. Dessa forma, a distribuição de probabilidades considerada na grande maioria das vezes é a distribuição Gaussiana. No entanto, outras distribuições podem ser consideradas.

Se for considerada a distribuição Gaussiana para a resposta da variável aleatória Y , supondo que a média é μ e a variância é constante, σ^2 , o modelo pode ser representado por

$$\begin{aligned} Y_i|x_i &\sim N(\mu_i, \sigma^2) \\ \mu_i &= \eta(x_i, \boldsymbol{\theta}). \end{aligned} \tag{7}$$

em que σ^2 e $\boldsymbol{\theta}$ são os parâmetros do modelo a serem estimados e η é a função não linear.

Um número crescente de pesquisadores compartilham que as relações entre variáveis biológicas são melhores descritas por funções não lineares. Processos de crescimento, decaimento, nascimento, mortalidade, abundância, competição e produção raramente são relacionadas linearmente às variáveis explicativas (SCHABENBERGER; PIERCE, 2002). E adicionalmente, em tais modelos, os parâmetros têm interpretações práticas (GRAYBILL; IYER, 2004).

Uma classe de modelos de regressão não lineares que têm proporcionado grande motivação para o desenvolvimento estatístico destes são os modelos de crescimento.

Segundo Seber e Wild (2003), a análise dos dados sobre crescimento é importante em muitas áreas de pesquisa. Na biologia, o interesse está na descrição do crescimento de animais e plantas na tentativa de compreender seus mecanismos essenci-

ais. Em química o interesse está na formulação do produto de uma reação química ao longo do tempo. Na agricultura, são óbvias as vantagens econômicas e administrativas em conhecer como a produção cresce, o quão rápido ela cresce, e como esses fatores respondem a condições ou tratamentos ambientais. O crescimento infantil é de interesse na medicina, assim como o crescimento de tumores e os efeitos de tratamentos sobre tais crescimentos. Cientistas sociais têm interesse em crescimento da população, no fornecimento de alimentos e na demanda de energia.

Na modelagem de dados de crescimento, deseja-se obter informações sobre a interpretação física dos parâmetros, a fim de construir um modelo padrão para as observações em estudo.

Para modelar esse tipo de comportamento, a literatura apresenta vários modelos de crescimento, ver Seber e Wild (2003). A seguir, são apresentados alguns dos principais modelos de crescimento.

2.4.1 Modelo de regressão não linear logístico

O modelo de regressão não linear logístico, também chamado de curva de crescimento logística, foi desenvolvido pelo matemático belga Verhulst (1838), como uma alternativa ao modelo de crescimento exponencial. Ele sugeriu que a taxa relativa de crescimento de η , em ordem a x , decresce linearmente com η . Ou seja,

$$\frac{\eta'(x)}{\eta(x)} = \frac{\beta_0 - \eta(x)}{s\beta_0}, \quad (8)$$

assim, para $\eta(x) \approx 0$, a taxa relativa de crescimento de η é aproximadamente constante ($1/s$) e a logística tem um crescimento aproximadamente exponencial. À medida que η cresce, a sua taxa relativa de crescimento vai tornando-se mais lenta. Para $\eta(x) \approx \beta_0$, a taxa relativa de crescimento de η é aproximadamente nula e a logística tende a estabilizar (deixar de crescer).

Assim, uma solução da equação diferencial 8 é a função não linear logística dada por

$$\eta(x; \boldsymbol{\theta}) = \frac{\beta_0}{1 + \exp\{-(x - x_I)/s\}}, \quad (9)$$

em que $\boldsymbol{\theta} = [\beta_0, x_I, s]^T$ são os parâmetros do modelo.

A curva de crescimento logística possui propriedades relevantes a respeito de pontos característicos, comportamento e padrões específicos. Assim, da equação (9), tais propriedades são definidas a seguir:

1. o parâmetro β_0 é uma assíntota horizontal à direita, pois

$$\lim_{x \rightarrow \infty} \eta(x; \boldsymbol{\theta}) = \beta_0,$$

β_0 representa a capacidade de sustentação do meio.

2. 0 é uma assíntota horizontal à esquerda, pois

$$\lim_{x \rightarrow -\infty} \eta(x; \boldsymbol{\theta}) = 0.$$

3. Como consequência dos itens 1. e 2., $\eta(x; \boldsymbol{\theta})$ tem imagem o intervalo $]0, \beta_0[$.

4. x_I é ponto de inflexão de $\eta(x; \boldsymbol{\theta})$, pois

$\eta(x; \boldsymbol{\theta})$ tem concavidade voltada para cima se $x < x_I$ e concavidade voltada para baixo se $x > x_I$.

5. $\eta(x; \boldsymbol{\theta})$ tem simetria em torno do seu ponto de inflexão:

$$\beta_0 - \eta(x_I - \delta) = \eta(x_I + \delta) - 0, \quad \forall \delta \geq 0.$$

E assim, $\eta(x_I) = \beta_0/2$.

6. $\eta(x; \boldsymbol{\theta})$ é uma função monótona crescente, pois

$$\eta(x_1) < \eta(x_2), \quad \forall x_1 < x_2,$$

o que implica em $\eta'(x) > 0$, ou seja, a função é crescente.

7. $\eta(x; \boldsymbol{\theta})$ possui ponto de fração q , pois, das propriedades 3 e 6, $\eta(x; \boldsymbol{\theta})$ é uma função monótona de contra domínio limitado, ou seja, $0 \leq \eta(x; \boldsymbol{\theta}) \leq \beta_0$, então, o valor x para o qual $\eta(x; \boldsymbol{\theta}) = q\beta_0$, $0 < q < 1$, é um ponto de fração q (PF_q). Assim, x_I é um ponto de fração 0,5 ($PF_{0,5}$).

8. $\eta'(x)$ tem ponto de máximo $x = x_I$, com valor $\eta'(x_I) = \beta_0/4s$.

9. $\eta(x; \boldsymbol{\theta})$ tem como padrão específico de comportamento uma sigmóide, ou seja, a curva tem formato de S .

Assim, para a parametrização da equação (9), verifica-se que x_I , s e β_0 têm unidade de medida X , X^{-1} e Y , respectivamente. Como já foi visto, o parâmetro x_I é o ponto de inflexão (PI), ponto de simetria (PS) e $PF_{0,5}$, portanto de tripla interpretação. Para s , verificou-se que tem relação com a taxa relativa de crescimento máxima. Por fim, β_0 é a assíntota superior.

O modelo de regressão não linear logístico é muito utilizado para descrever o crescimento de populações, de frutos, plantas e animais. Prado et. al. (2013); Cunha (2011); Fernandes et. al (2013) utilizaram o modelo para estudar o crescimento de frutos. Para o estudar o crescimento de frangos, Veloso et. al. (2015) e Moraes et. al. (2015) também utilizaram o modelo de regressão não linear logístico.

2.4.2 Modelo de regressão não linear de Gompertz

O modelo de regressão de Gompertz, conhecido também como lei de Gompertz, foi desenvolvido, originalmente, pelo matemático britânico Benjamin Gompertz para estimar a mortalidade humana (GOMPERTZ, 1825). Winsor (1932), apresenta uma descrição inicial da utilização desta equação para descrever o processo de crescimento.

O modelo de Gompertz é uma outra alternativa ao modelo de crescimento exponencial, e tem como equação diferencial

$$\frac{\eta'(x)}{\eta(x)} = s \exp\{-s(x - x_I)\}, \quad (10)$$

que corresponde a dizer que a taxa relativa de crescimento de η , em ordem a x , decresce exponencialmente com o aumento de x .

Uma solução da equação diferencial (10) é a função não linear de Gompertz, dada por

$$\eta(x; \boldsymbol{\theta}) = \beta_0 \exp\{-\exp\{-s(x - x_I)\}\}, \quad (11)$$

em que $\boldsymbol{\theta} = [\beta_0, x_I, s]^T$ são os parâmetros do modelo.

O modelo de Gompertz é muito parecido com o logístico. As propriedades são as mesmas, com exceção de que:

1. $\eta(x; \boldsymbol{\theta})$ não tem simetria em torno do seu ponto de inflexão $x = x_I$, em que $\eta(x_I) = \beta_0 / \exp\{1\} < \beta_0/2$, ou seja, η atinge menos da metade da capacidade de sustentação. E assim, x_I é um ponto de fração $1/\exp\{1\}$ ($PF_{1/\exp\{1\}}$).
2. $\eta'(x)$ tem ponto de máximo $x = x_I$, com valor $\eta'(x_I) = s\beta_0/\exp\{1\}$.

Logo, as interpretações dos parâmetros do modelo de Gompertz são as mesmas do modelo logístico. Ou seja, o parâmetro β_0 é a capacidade de sustentação do meio e a curva é também uma sigmóide, que devido as propriedades 1 e 2 relacionadas aos parâmetros x_I e s , a torna mais flexível que a curva logística. Para estudar o crescimento de animais, Mota et. al. (2015) e Falcão et. al. (2015) utilizaram o modelo de regressão não linear de Gompertz.

2.5 Reparametrização dos modelos não lineares

Existem várias parametrizações para cada modelo de regressão não linear. Assim, existe mais que uma representação para o modelo não linear logístico, de Gompertz, e outros. As várias representações para um dado modelo específico, trata-se de um mesmo modelo, ou seja, descrevem a mesma curva, porém, com seus parâmetros reparametrizados e por isso se diz que são parametrizações de um mesmo modelo.

De acordo com Zeviani (2013), é possível reparametrizar um modelo $\eta(x, \boldsymbol{\theta})$ para um parâmetro de interesse ϑ caso ele possa ser escrito como função dos parâmetros do modelo, $\vartheta = g(\boldsymbol{\theta})$. O procedimento é simples e envolve três etapas:

1. Escrever o parâmetro de interesse como função dos elementos de $\boldsymbol{\theta}$, ou seja, $\vartheta = g(\boldsymbol{\theta})$;
2. Escolher um dos elementos de $\boldsymbol{\theta} = (\theta_i, \boldsymbol{\theta}_{-i})$ para ser colocado em função de ϑ de tal forma a obter $\theta_i = h(\boldsymbol{\theta}_{-i}, \vartheta)$;
3. Substituir toda ocorrência de θ_i em $\eta(x, \boldsymbol{\theta})$ pela expressão obtida no passo anterior, $h(\boldsymbol{\theta}_{-i}, \vartheta)$, fazendo as simplificações convenientes. Assim o modelo de regressão não linear

$$\eta(x, \boldsymbol{\theta}_{-i}, \vartheta)$$

terá ϑ como elemento do vetor de parâmetros $\boldsymbol{\theta}_\vartheta = (\boldsymbol{\theta}_{-i}, \vartheta)$.

É necessário que a função g seja estritamente monótona e diferenciável sobre o domínio relacionado à θ_i e dessa forma a função h é a inversa de g em θ_i .

Recomenda-se no passo 2 priorizar àquele elemento de $\boldsymbol{\theta}$ com menor significado ou para o qual a obtenção de h seja menos complicada. Em algumas situações pode ser que h não tenha solução analítica, por exemplo, quando é impossível isolar θ_i , nesse caso, faz-se necessário aplicação de métodos numéricos. O procedimento também pode ser usado para mais de um parâmetro de interesse, basta repeti-lo até que todos $(\vartheta_1, \dots, \vartheta_k, k \geq p)$ sejam incorporados.

Uma etapa importante na reparametrização é fazer o estudo de dimensionalidade para determinar as unidades de medida dos parâmetros bem como seus espaços paramétricos. Tais informações são relevantes para descrição do modelo, interpretação dos parâmetros e conversão de valores entre escalas de medida.

2.6 Métodos de estimação dos modelos não lineares

2.6.1 Verossimilhança

Como já foi dito e exemplificado, os modelos de regressão não lineares geralmente são utilizados para respostas contínuas. Assim, a distribuição de probabilidade considerada na maioria das vezes é a distribuição Gaussiana. No entanto, outras distribuições podem ser consideradas.

Considerando o modelo representado pela equação (7), temos que a função de verossimilhança do modelo é dada por

$$L(\boldsymbol{\theta}, \sigma^2) = \prod_{i=1}^n \phi(y_i, \eta(x_i, \boldsymbol{\theta}), \sigma^2),$$

em que ϕ representa a função densidade da distribuição Normal. A estimativa de máxima verossimilhança são os valores $(\hat{\boldsymbol{\theta}}, \hat{\sigma}^2)$ que tornam máximo valor de L . Para estimação, é conveniente trabalhar com o logaritmo da função de verossimilhança

$$\ell(\boldsymbol{\theta}, \sigma^2) = \log [L(\boldsymbol{\theta}, \sigma^2)] .$$

Um intervalo de confiança para $\theta_i, i = 1, \dots, p$, baseado na função deviance $D(\theta_i)$ é definido por θ_i : $D(\theta_i) \leq c$ em que $D(\theta_i) = -2 [\ell(\theta_i) - \ell(\hat{\boldsymbol{\theta}})]$. A quantidade $\ell(\theta_i)$

representa o valor maximizado da log-verossimilhança com relação aos demais parâmetros para um valor fixo de θ_i enquanto que $\ell(\hat{\theta})$ é o valor máximo da log-verossimilhança com relação à todos os parâmetros. O valor c geralmente é definido de forma que o intervalo possua uma interpretação frequentista e se baseia na distribuição amostral da função deviance que é a distribuição qui-quadrado. Por exemplo, para obter um intervalo de confiança com cobertura nominal de 95% o valor de c é $\chi_1^2 = 3,84$.

Para encontrar o limites do intervalo de confiança para θ_i é necessário encontrar as raízes da equação $D(\theta_i) - c = 0$. No caso de modelos de regressão não linear, não é possível obter expressões fechadas para os limites do intervalo, portanto métodos numéricos são aplicados.

Os intervalos obtidos dessa forma apresentam alguma assimetria devido à verossimilhança ser uma função não simétrica em θ_i . O grau de assimetria diminui com o tamanho da amostra. O gráfico de $D(\theta_i)$ em função de θ_i , chamado de gráfico de perfil de log-verossimilhança para o parâmetro θ_i , permite visualizar tanto os limites do intervalo com relação ao valor estabelecido de c quanto o grau de simetria da função. Para Ritz e Streibig (2008), o gráfico de perfil é útil para acessar a extensão da curvatura do modelo na direção do parâmetro θ_i . Mais detalhes sobre o perfil de verossimilhança estão disponíveis em Bates e Watts (1988) e Venables e Ripley (2002).

Clarke (1987) aponta que a maior dificuldade no uso da função de verossimilhança para obter intervalos de confiança está na quantidade de trabalho envolvida para obter os seus limites. Tipicamente, isso requer sucessivas maximizações da verossimilhança com relação aos demais parâmetros para valores de θ_i numa sequência qualquer (MATTHEWS, 1988).

2.6.2 Aproximação quadrática da verossimilhança

Dada a restrição computacional para obtenção de intervalos baseados em verossimilhança, é comum simplificar a obtenção do intervalo de confiança para θ_i por meio de uma aproximação quadrática da função de log-verossimilhança ao redor de $\hat{\theta}$. Procedendo dessa forma temos

$$\begin{aligned} D(\theta_i) &= -2[\ell(\theta_i) - \ell(\hat{\theta}_i)] \\ &\approx -2 \left[\ell(\hat{\theta}_i) + (\theta_i - \hat{\theta}_i)\ell'(\hat{\theta}_i) + \frac{1}{2}(\theta_i - \hat{\theta}_i)^2\ell''(\hat{\theta}_i) - \ell(\hat{\theta}_i) \right]. \end{aligned}$$

Uma vez que $\hat{\theta}_i$ é estimador de máxima log-verossimilhança, $\ell'(\hat{\theta}_i) = 0$. Assim define-se o intervalo de confiança aproximado por meio de

$$D(\theta_i) = -(\theta_i - \hat{\theta}_i)^2 \ell''(\hat{\theta}_i) \leq c,$$

com solução do tipo

$$\hat{\theta}_i \pm \sqrt{\frac{c}{\ell''(\hat{\theta}_i)}}.$$

Quando a distribuição considerada para a variável dependente é normal com variância constante, usar a aproximação da função de log-verossimilhança é equivalente a usar uma aproximação da função η para obter os intervalos de confiança. Em modelos lineares, a soma de quadrado dos resíduos

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

é quadrática e seus contornos de valor constante são elipsoides. Em modelos não lineares, se aproximarmos a função η por uma série de Taylor ao redor de $\hat{\boldsymbol{\theta}}$, a soma de quadrados resultantes da aproximação será quadrática. Isso resulta em intervalos individuais simétricos centrados em $\hat{\theta}_i$.

Segundo Seber e Wild (2003), fazendo uso de uma aproximação linear da função η e para uma amostra suficientemente grande, $\hat{\boldsymbol{\theta}}$ apresentará distribuição amostral normal, ou seja,

$$\hat{\boldsymbol{\theta}} \sim \text{Normal}(\boldsymbol{\theta}, \sigma^2(\mathbf{F}^\top \mathbf{F})^{-1}), \quad (12)$$

em que $\mathbf{F} = \frac{\partial \eta(\mathbf{x}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}^\top}$ representa a matriz de delineamento dada pela aproximação linear da função η avaliada em $\hat{\boldsymbol{\theta}}$. Um intervalo de confiança de cobertura $1 - \alpha$ para θ_i é então obtido por

$$\hat{\theta}_i \pm t_{n-p} \sqrt{\hat{\sigma}_{ii}},$$

em que t_{n-p} é o quantil $1 - \alpha/2$ da distribuição t com $n - p$ graus de liberdade, $\hat{\sigma}_{ii}$ representa a estimativa de variância de $\hat{\theta}_i$ e é o i -ésimo elemento da diagonal da matriz

de variância-covariância $\hat{\sigma}^2(\mathbf{F}^\top \mathbf{F})^{-1}$, sendo $\hat{\sigma}^2 = \frac{SQR(\hat{\theta})}{n-p}$. Intervalos obtidos dessa forma, do tipo mais ou menos, são chamados de intervalos de Wald.

A vantagem do procedimento de Wald é que não requer a avaliação consecutiva de nenhuma função como é o intervalo baseado na verossimilhança, apenas do cálculo da matriz $\hat{\sigma}^2(\mathbf{F}^\top \mathbf{F})^{-1}$ após obter as estimativas dos parâmetros.

2.7 Seleção de modelos

2.7.1 Medidas de curvatura

Os procedimentos de inferência em modelos de regressão não linear fazem uso de argumentos assintóticos e frequente analogia com modelos de regressão linear (DRAPER; SMITH, 1988; SEBER; WILD, 2003). Em outras palavras, as inferências são baseadas numa aproximação linear do modelo não linear.

Para avaliar a qualidade de aproximação linear, Bates e Watts (1980) propuseram medidas de curvatura que se baseiam na curvatura da função não linear $\eta(x, \theta)$ por meio da magnitude da segunda derivada $\partial^2 \eta / \partial \theta_j \partial \theta_k$, com $j, k = 1, \dots, p$, p é a dimensão de θ , n é o número de observações. Se o valor absoluto da segunda derivada é pequeno, então o modelo é aproximadamente linear e as suposições assintóticas se tornam aceitáveis. Seja $\mathbf{u} = (u_1, \dots, u_n)^\top$ um vetor que representa uma direção de interesse no espaço vetorial a partir de $\hat{\theta}$ e defina

$$A_{jk} = \sum_{i=1}^n u_i \frac{\partial^2 \eta(\mathbf{x}_i, \hat{\theta})}{\partial \theta_j \partial \theta_k}.$$

A matriz $\mathbf{A}_{(p \times p)} = \{A_{jk}\}$ representa o valor da segunda derivada na direção $\mathbf{u}$. Para obter medidas de curvatura é necessário padronizar as derivadas pelo valor do gradiente na mesma direção. Seja $\mathbf{X} = \mathbf{QR}$ a decomposição QR da matriz $\mathbf{X}_{(n \times p)} = \left. \frac{\partial \eta(\mathbf{x}_i, \hat{\theta})}{\partial \theta_j} \right|_{\theta = \hat{\theta}}$, e seja

$$\mathbf{C}_{(p \times p)} = (\mathbf{R}^{-1})^\top \mathbf{A} \mathbf{R}^{-1}.$$

Bates e Watts (1980) usam $\mathbf{C}$ para definirem as medias de curvatura relativas à não linearidade. Se $\mathbf{u}$ pertence ao espaço vetorial gerado pelas colunas de $\mathbf{X}$ ($\mathbf{u} = \mathbf{X}\theta$ para um dado θ), então $\mathbf{C}$ representa a curvatura devido ao efeito de parâmetros.

Se $\mathbf{X}^\top \mathbf{u} = \mathbf{0}$ então $\mathbf{C}$ representa a curvatura intrínseca do modelo. Pode-se mostrar que a curvatura intrínseca do modelo depende apenas da forma da superfície esperada, gerada pelos valores preditos pelo modelo $\eta(\mathbf{x}, \boldsymbol{\theta})$, à medida que $\boldsymbol{\theta}$ varia e que por isso é invariante em relação à parametrização adotada (SEBER; WILD, 2003).

2.7.2 Critério AIC e BIC

Os critérios de Akaike (AIC) e Bayesiano (BIC) são baseados no valor do logaritmo natural da função de verossimilhança do modelo, e dependem do número de observações e do número de parâmetros do modelo.

Segundo Akaike (1974), o critério de AIC é calculado por

$$AIC = -2l(\hat{\boldsymbol{\theta}}) + 2p,$$

em que p é o número total de parâmetros do modelo.

O critério BIC é bastante semelhante, penalizando com mais severidade os modelos que apresentam maior número de parâmetros, segundo a expressão

$$BIC = -2l(\hat{\boldsymbol{\theta}}) + p \log(n),$$

em que n é o número de observações utilizadas na estimação dos parâmetros do modelo, e p como definido para AIC.

Valores baixos para AIC e BIC são preferidos, indicando o modelo a ser escolhido dentro de uma comparação de modelos. Valores altos de verossimilhança implicam em baixos valores de AIC e BIC. Ambos critérios penalizam modelos com grande número de parâmetros, optando-se por modelos mais parcimoniosos. Nem sempre os dois critérios indicam um mesmo modelo, cabendo ao pesquisador decidir que critério seguir.

3 MODELOS DA SOMA DE REGRESSÕES LINEARES PONDERADAS POR FUNÇÕES DE DISTRIBUIÇÃO ACUMULADA

3.1 Modelo geral

Consideremos um modelo de regressão dado pela soma ponderada de regressões lineares em que os pesos associados à cada regressão são funções matemáticas que possuem as propriedades de uma função de distribuição acumulada. Assim temos que:

$$\eta(x; \boldsymbol{\theta}) = \sum_{j=1}^m f_j(x; \boldsymbol{\beta}_j) \pi_j(x; \boldsymbol{\alpha}_j) \quad (13)$$

em que $j = 1, \dots, m$ denominaremos de regimes; $f_j = \beta_{j0} + \beta_{j1}x + \dots + \beta_{jp_j}x^{p_j}$ é a j -ésima regressão linear de grau p_j , $p_j = 0, 1, 2, \dots, n$ e vetor de parâmetros $\boldsymbol{\beta}_j = [\beta_{j0}, \beta_{j1}, \dots, \beta_{jp_j}]^T$; π_j é a j -ésima função de distribuição acumulada com vetor de parâmetros $\boldsymbol{\alpha}_j$. Então η é um modelo de regressão não linear com vetor de parâmetros $\boldsymbol{\theta} = [\boldsymbol{\beta}_j, \boldsymbol{\alpha}_j]^T$.

Prova: Seja f_j a j -ésima regressão linear de grau p_j da soma ponderada da equação 13. Assim, β_{j0} pertence a $\boldsymbol{\beta}_j$ para qualquer que seja o valor de p_j . Daí,

$$\begin{aligned} \frac{\partial \eta(x; \boldsymbol{\theta})}{\partial \beta_{j0}} &= \frac{\partial f_j(x; \boldsymbol{\beta}_j)}{\partial \beta_{j0}} \pi_j(x; \boldsymbol{\alpha}_j) \\ \frac{\partial \eta(x; \boldsymbol{\theta})}{\partial \beta_{j0}} &= \pi_j(x; \boldsymbol{\alpha}_j), \end{aligned}$$

portanto, η é um modelo de regressão não linear resultante da soma ponderada de regressões lineares em que cada peso associado é uma função de distribuição acumulada (NLFDA). Com isso, pode-se criar vários modelos de regressão não linear, em que as características de cada modelo serão idênticas às da função distribuição acumulada a ser utilizada.

3.2 Caracterização dos modelos de regressão de dois regimes

Consideremos um modelo de regressão não linear da equação (13), com $j = 2$. Se as duas regressões lineares são dadas por $f_1(x) = \beta_{10}$ e $f_2(x) = \beta_{20}$, ou seja, duas constantes, em que os pesos π_1 e π_2 , possuem a relação $\pi_1(x) = 1 - \pi_2(x)$, então,

$$\eta(x; \boldsymbol{\theta}) = \beta_{10}[1 - \pi_2(x)] + \beta_{20}\pi_2(x)$$

logo,

$$\eta(x; \boldsymbol{\theta}) = \beta_{10} + (\beta_{20} - \beta_{10})\pi_2(x) \quad (14)$$

Modelo de regressão não linear da Logística: Seja $\pi_2(x) = 1/(1+\exp\{-(x-x_I)/s\})$, ou seja uma FDA da logística, então, da equação (14), tem-se

$$\eta(x; \boldsymbol{\theta}) = \beta_{10} + \frac{\beta_{20} - \beta_{10}}{1 + \exp\{-(x - x_I)/s\}} \quad (15)$$

em que $\boldsymbol{\theta} = [\beta_{10}, \beta_{20}, x_I, s]^T$, que é exatamente o modelo de regressão não linear logístico com duas assíntotas horizontais.

Um caso particular desse modelo é se $\beta_{10} = 0$, assim

$$\eta(x; \boldsymbol{\theta}) = \frac{\beta_{20}}{1 + \exp\{-(x - x_I)/s\}} \quad (16)$$

em que $\boldsymbol{\theta} = [\beta_{20}, x_I, s]^T$, que é o mesmo modelo de regressão não linear Logístico de apenas uma assíntota, de equação (9).

Modelo de regressão não linear Gumbel: Considerando agora, $\pi_2(x)$, uma FDA da Gumbell, com parametrização $\pi_2(x) = \exp\{-\exp\{-s(x - x_I)\}\}$. Assim, de (14), tem-se

$$\eta(x; \boldsymbol{\theta}) = \beta_{10} + (\beta_{20} - \beta_{10}) \exp\{-\exp\{-s(x - x_I)\}\} \quad (17)$$

em que $\boldsymbol{\theta} = [\beta_{10}, \beta_{20}, x_I, s]^T$, que é exatamente o modelo de regressão não linear de Gompertz com duas assíntotas horizontais.

Se $\beta_{10} = 0$, tem-se

$$\eta(x; \boldsymbol{\theta}) = \beta_{20} \exp\{-\exp\{-s(x - x_I)\}\} \quad (18)$$

em que $\boldsymbol{\theta} = [\beta_{20}, x_I, s]^T$, que é o mesmo modelo de regressão não linear de Gompertz de apenas uma assíntota, de equação (11).

Modelo de regressão não linear Cauchy seno hiperbólico exponencial: Como vimos na seção 2, a FDA da Exp-SC é bastante flexível assumindo várias formas depen-

dendo dos valores de seus parâmetros. Assim, seja

$\pi_2(x) = \{1/2 + [1/\pi] \arctan[\nu \sinh((x - x_I)/s)]\}^\tau$, ou seja, uma FDA da distribuição Exp-SC, logo, de (14), tem-se

$$\eta(x; \boldsymbol{\theta}) = \beta_{10} + (\beta_{20} - \beta_{10}) \left\{ \frac{1}{2} + \frac{1}{\pi} \arctan \left[\nu \sinh \left(\frac{x - x_I}{s} \right) \right] \right\}^\tau \quad (19)$$

em que $\boldsymbol{\theta} = [\beta_{10}, \beta_{20}, x_I, s, \nu, \tau]^T$ é o vetor de parâmetros do modelo.

Assumindo-se $\beta_{10} = 0$, tem-se

$$\eta(x; \boldsymbol{\theta}) = \beta_{20} \left\{ \frac{1}{2} + \frac{1}{\pi} \arctan \left[\nu \sinh \left(\frac{x - x_I}{s} \right) \right] \right\}^\tau. \quad (20)$$

em que $\boldsymbol{\theta} = [\beta_{20}, x_I, s, \nu, \tau]^T$.

Denominou-se (19) e (20) de modelos de regressão não lineares Exp-SC de duas e uma assíntota horizontal, respectivamente. As interpretações dos parâmetros x_I , s , ν e τ são as mesmas da FDA Exp-SC, ou seja, a curva característica do modelo de regressão não linear é a mesma da FDA, mudando apenas a imagem da função, pois β_{10} e β_{20} são as assíntotas horizontais.

Portanto, temos que os modelos de regressão não lineares (14) possuem as mesmas características de sua FDA correspondente em relação à sua forma. Com isso, pode-se criar vários modelos de regressão não lineares que herdaram as características da FDA utilizada.

3.2.1 Reparametrização no ponto de fração (PF_q)

De acordo com Zeviani (2013), é possível reparametrizar um modelo não linear $\eta(x, \boldsymbol{\theta})$ de forma que se obtenha o ponto de fração x_q caso ele possa ser escrito como função dos parâmetros do modelo, $x_q = g(\boldsymbol{\theta})$. Assim, podemos reparametrizar o modelo para que qualquer fração do total seja um de seus parâmetros.

Para o modelo de equação (14), tem-se que o valor x_q da fração q de β_{20} é

$$\begin{aligned} q\beta_{20} &= \beta_{10} + (\beta_{20} - \beta_{10})\pi_2(x_q), \quad \text{então} \\ x_q &= \pi_2^{-1} \left(\frac{q\beta_{20} - \beta_{10}}{\beta_{20} - \beta_{10}} \right), \quad \text{logo} \\ x_q &= Q(q^*) \end{aligned} \quad (21)$$

em que, Q é a função quantílica ou seja, a inversa da FDA, e $q^* = (q\beta_{20} - \beta_{10})/(\beta_{20} - \beta_{10})$.

Para os casos em que $\beta_{10} = 0$, temos $q^* = q$, logo, o ponto x_q é expresso apenas em função dos parâmetros da FDA do modelo. Logo, uma reparametrização do modelo de equação (14) em um ponto de fração q de β_{20} se dá exclusivamente pela reparametrização da FDA no ponto de fração escolhido.

4 MATERIAL E MÉTODOS

4.1 Tempos de resposta de um sistema eletrônico de pulverização

Os dados para obtenção do tempo de resposta de um sistema de controle eletrônico de pulverização são de Vieira (2013), onde realizou ensaios no Laboratório de Tecnologia de Aplicação, localizado no Departamento de Engenharia de Biosistemas da Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, no município de Piracicaba, São Paulo, Brasil.

Uma bancada constituída por: controlador eletrônico de pulverização, simulador de velocidade, base de aquisição de dados e sistema hidráulico foi projetada para realização dos ensaios visando simular o experimento em campo.

O sistema eletrônico foi composto de controlador eletrônico Vcom 8.4[®], Verion Agricultura Inc. (Brasil), simulador de velocidade MFG-4221 20MHz DDS Minipa[®] Inc. (EUA), válvulas elétricas reguladora e de corte automático de seção Geoline, Geoline[®] Inc. (Itália), transdutor de fluxo com turbina axial removível de marca Polmac[®], Polmac Inc. (Itália).

As pontas de pulverização utilizadas foram Teejet, Spraying System[®], do modelo XR com ângulo de aspersão de 110°, ambas produzindo jato leque plano padrão de uso amplificado: XR 110 02 (baixa vazão), XR 110 04 (média vazão) e XR 110 06 (alta vazão) (SPRAYING SYSTEM, 2011).

Com o intuito de submeter o sistema às mais próximas condições operacionais encontradas no campo, foram realizadas simulações para as três diferentes pontas de pulverização fazendo-se mudanças na velocidade e nas taxas de aplicação (vazão). A variação da pressão (kPa) antes e após a mudança realizada em função do tempo (s) foram coletadas. Assim, o tempo de resposta, t_r , é igual a diferença do valor do tempo em que a pressão volta a estabilizar, t_e , pelo tempo em que realizou-se a mudança, t_m . Detalhes são encontrados em Vieira (2013).

Nesse trabalho, determinou-se os tempos de resposta apenas para as mudanças de velocidade, sendo que as alterações foram feitas respeitando as especificações em manual da ponta para velocidade mínima e máxima que cada ponta de pulverização utilizada suporta para seu melhor desempenho. A taxa de aplicação foi fixada em 120 litros por hectare (L/ha).

Nas Figuras 4, 5 e 6 tem-se o diagrama de dispersão dos valores observados da pressão em Quilopascal (kPa) em função do tempo, em segundos (s), para as pontas de pulverização XR 110 02, XR 110 04 e XR 110 06, de cada mudança de velocidade, em que utilizou-se os intervalos mínimo e máximo de cada ponta que são de 4 a 8, 10 a 18 e 16 a 25 km/h, respectivamente. O tempo exato de mudança foi fixo, $t_m = 2s$.

Figura 4 – Pressão (kPa) em função do tempo (s), para a ponta de pulverização XR 110 02 (baixa vazão), de cada mudança de velocidade

Figura 5 – Pressão (kPa) em função do tempo (s), para a ponta de pulverização XR 110 04 (média vazão), de cada mudança de velocidade

Figura 6 – Pressão (kPa) em função do tempo (s), para a ponta de pulverização XR 110 06 (alta vazão), de cada mudança de velocidade

Das figuras apresentadas, observa-se que as mudanças em que ocorre o aumento da velocidade (ascendente), tem-se uma relação monótona não decrescente. Para mudanças que há decréscimo da velocidade (descendente), tem-se uma relação monótona não crescente. Observa-se também um padrão sigmoidal de duas assíntotas, em que a transição ocorre de forma simétrica ou assimétrica, dependendo da mudança realizada. Assim, as funções matemáticas que devem atender essas características dentre os vários modelos não lineares são os logístico e de Gompertz, de duas assíntotas horizontais, de equações (15) e (17), respectivamente. A assíntota de menor valor denominamos assíntota inferior e a de maior valor assíntota superior.

Para incluir o tempo de resposta, t_r , como um dos parâmetros, dos dois modelos a ser ajustados, foi feita uma reparametrização, considerando as regras de Zeviani (2013), em um ponto de fração, q , da assíntota β_{20} , ou seja, da assíntota superior na ascendente e da assíntota inferior na descendente. Denominamos esse ponto de tempo de estabilização, t_e , tempo no qual a pressão volta estabilizar após a mudança realizada. Mas, como o tempo de resposta, t_r , é igual a diferença entre o tempo de estabilização, t_e , e o ponto em que houve a mudança, $t_m = 2s$, então $t_e = t_r + 2$, e assim inclui-se o novo parâmetro no modelo.

Para o modelo de regressão não linear logístico (15), temos que a função inversa da FDA da logística é dada por

$$Q(x) = x_I - s \log \left(\frac{1}{x} - 1 \right)$$

e assim, da equação (21), temos

$$t_e = x_I - s \log \left(\frac{1 - q^*}{q^*} \right)$$

logo,

$$t_r = x_I - s \log \left(\frac{1 - q^*}{q^*} \right) - 2.$$

Reescrevendo a equação acima de forma a isolar o parâmetro s , tem-se

$$s = \frac{x_I - t_r - 2}{\log \left(\frac{1 - q^*}{q^*} \right)},$$

por fim, substituindo s no modelo (15) e, após algumas simplificações, tem-se

$$\eta(x; \boldsymbol{\theta}) = \beta_{10} + \frac{\beta_{20} - \beta_{10}}{\left(1 + \frac{1 - q^*}{q^*} \right)^{\frac{x_I - x}{x_I - 2 - t_r}}} \quad (22)$$

em que $\boldsymbol{\theta} = [\beta_{10}, \beta_{20}, x_I, t_r]^T$ e $q^* = (q\beta_{20} - \beta_{10})/(\beta_{20} - \beta_{10})$, é o modelo de regressão não linear logístico em que o tempo de resposta é um dos seus parâmetros.

Para o modelo de regressão não linear de Gompertz (17), a função inversa da FDA de Gumbel é dada por

$$Q(x) = x_I - s \log(-\log(x))$$

assim,

$$t_e = x_I - s \log(-\log(q^*))$$

então, de modo análogo ao modelo anterior, temos

$$s = \frac{x_I - t_r - 2}{\log(-\log(q^*))}$$

logo,

$$\eta(x; \boldsymbol{\theta}) = \beta_{10} + (\beta_{20} - \beta_{10}) \exp \left\{ \frac{-1}{-\log(q^*)^{\left(\frac{x - x_I}{x_I - 2 - t_r} \right)}} \right\} \quad (23)$$

em que $\boldsymbol{\theta} = [\beta_{10}, \beta_{20}, x_I, t_r]^T$ e $q^* = (q\beta_{20} - \beta_{10})/(\beta_{20} - \beta_{10})$, é o modelo de regressão não linear de Gompertz com o parâmetro t_r incluso no modelo.

Baseado num critério descritivo para encontrar o tempo de resposta, adotado por Gadanha Junior (2000), considerou-se as frações $q = 0,95$, da assíntota superior, e $q = 1,05$ da assíntota inferior, como aceitável para representar o ponto de estabilização, t_e , nas mudanças de velocidade. Assim, os modelos a ser ajustados para encontrar os tempos de resposta foram:

$$\eta(x; \boldsymbol{\theta}) = \beta_{10} + \frac{\beta_{20} - \beta_{10}}{\left(1 + \frac{1-q^*}{q^*}\right)^{\frac{x_I - x}{x_I - 2 - t_r}}} \quad (24)$$

$$\eta(x; \boldsymbol{\theta}) = \beta_{10} + (\beta_{20} - \beta_{10}) \exp \left\{ \frac{-1}{-\log(q^*) \left(\frac{x - x_I}{x_I - 2 - t_r}\right)} \right\} \quad (25)$$

em que $\boldsymbol{\theta} = [\beta_{10}, \beta_{20}, x_I, t_r]^T$ e $q^* = (q\beta_{20} - \beta_{10})/(\beta_{20} - \beta_{10})$, $q = (0,95; 1,05)$.

Para verificar a adequação do ajuste dos modelos aos dados foram feitas as análises dos resíduos e posteriormente calculadas as medidas de curvatura de Bates e Watts (1980). Tais medidas também foram utilizadas, juntamente com o critério de Akaike (1974) para seleção dos modelos. Inferências para o parâmetro tempo de resposta t_r foram obtidas por meio da verossimilhança e da distribuição assintótica dos estimadores. Utilizou-se o software estatístico R (2015) para realização desse estudo.

4.2 Dados simulados

Segundo Ratkowsky (1983), estudos via simulação são as melhores e mais diretas maneiras para estudar as propriedades amostrais dos estimadores de máxima verossimilhança. Assim, para obtenção de resultados mais abrangentes do modelo de equação (14), um estudo de simulação foi realizado, utilizando a metodologia de Monte Carlo, para avaliar as estimativas de máxima verossimilhança.

Para gerar os dados, utilizou-se a equação (20), com parâmetros, $\beta_{20} = 50$; $x_I = 50$; $s = 5$; $\nu = 0,1$ e $\tau = 0,6$ considerando erros com distribuição normal e identicamente distribuídos com média zero e variância constante. A variável independente x foi fixada num intervalo de 7 a 90, em que a quantidade de pontos amostrais, h , dentro desse intervalo multiplicado pelo número de repetições, r , constitui o tamanho amostral n . Considerou-se o número de repetições fixo $r = 3$, e pontos amostrais $h = 6, 15, 45, 90$,

logo, os quatro tamanhos amostrais utilizados para esse estudo foram $n = 18, 45, 135, 270$. O intervalo e os pontos amostrais da variável independente x foram escolhidos baseados em um estudo com ameixa, realizado por Cunha (2011).

Para cada amostra, n , o processo foi repetido 1000 vezes e calculou-se, para cada parâmetro, a estimativa média (EM), o viés definido como $E(\hat{\theta}) - \theta$, o erro quadrático médio (EQM), o coeficiente de assimetria (CA) e o coeficiente de curtose (K). Os valores observados para os verdadeiros parâmetros selecionados e para as estimativas médias (EMs), em função da variável independente x , foram apresentados. Por fim, os gráficos com as estimativas de densidade Kernel, para cada parâmetro, das 1000 amostragens de cada tamanho amostral foram feitos no intuito de apresentar graficamente as propriedades amostrais dos estimadores dos parâmetros. Também utilizou-se o software estatístico R (2015) para realização desse estudo.

5 RESULTADOS E DISCUSSÃO

5.1 Avaliação dos tempos de resposta

Os ajustes dos modelos aos dados convergiram para todas as mudanças de velocidade nas três pontas de pulverização, para os dois modelos utilizados, considerando o máximo de 50 iterações. Valores iniciais baseados na inspeção do diagrama de dispersão foram considerados (Figuras 4, 5 e 6). A análise gráfica dos resíduos para os modelos propostos está apresentada no Apêndice A.

Para a ponta de pulverização de baixa vazão XR 110 02 verificou-se que o tempo de resposta estimado para o modelo Gompertz foi maior que o correspondente no modelo logístico, para cada mudança de velocidade (Tabela 1).

Tabela 1 – Estimativa (Est.), limite inferior (LI) e limite superior (LS) para o intervalo de confiança assintótico (95%) para o parâmetro de interesse tempo de resposta, t_r , dos modelos Logístico e de Gompertz para a ponta de baixa vazão

Velocidade (km/h)	Modelo ajustado	Tempo de Resposta (t_r)		
		LI	LS	Est.
4 a 8	Logístico	6,01	6,36	6,19
	Gompertz	6,88	7,62	7,25
8 a 4	Logístico	5,32	6,05	5,69
	Gompertz	6,17	6,88	6,53

Ao avaliar à extensão da não linearidade pelas medidas de curvatura de Bates e Watts (1980), verificou-se que tanto a curvatura devido à parametrização (CP) quanto a intrínseca (CI) foram muito próximas para ambos os modelos para as duas mudanças de velocidade (Tabela 2). Para todos ajustes os valores da CI são aceitáveis, considerando 0,3 como ponto de corte (ROBINSON; HAMANN, 2011). Por essas medidas de curvatura e pelo AIC observa-se que para a mudança que ocorre o aumento da velocidade, o modelo logístico obteve melhor resultado em relação ao modelo de Gompertz. Já para a mudança no decréscimo da velocidade, os melhores resultados são obtidos pelo modelo de Gompertz.

Valores preditos para os modelos selecionados, de acordo com as medidas de não linearidade e pelo AIC, estão representados na Figura 7, juntamente com a banda de confiança (95%) e com destaque para a estimativa intervalar assintótica (95%) do tempo de resposta do sistema para a ponta de baixa vazão XR 110 02.

Tabela 2 – Medidas do critério de informação de Akaike (AIC) e de curvatura devido ao efeito de parametrização (CP) e intrínseca (CI) dos modelos logístico e de Gompertz para a ponta de baixa vazão

Velocidade (km/h)	Modelo	Curvatura		AIC
		CP	CI	
4 a 8	Logístico	0,0714	0,0373	354,17
	Gompertz	0,1006	0,0579	378,28
8 a 4	Logístico	0,1149	0,0530	443,42
	Gompertz	0,0721	0,0522	395,21

Figura 7 – Valores observados e preditos com banda de confiança assintótica (95%) para a pressão (kPa) em função do tempo (s) para cada mudança de velocidade. Linhas verticais identificam a estimativa intervalar assintótica (95%) do tempo de resposta do controlador eletrônico de pulverização para a ponta de baixa vazão XR 110 02

Observa-se, da Tabela 3, que os tempos de resposta são estatisticamente iguais, pois os intervalos de confiança estão sobrepostos, assim, tanto na ascendente quanto na descendente, o tempo em que o sistema gasta para voltar a estabilizar devido às mudanças de velocidade é o mesmo para a ponta de baixa vazão utilizada.

Tabela 3 – Estimativa (Est.), limite inferior (LI) e limite superior (LS) para o intervalo de confiança assintótico (95%) para o parâmetro de interesse tempo de resposta, t_r , dos modelos selecionados para a ponta de baixa vazão

Velocidade (km/h)	Modelo	Tempo de Resposta (t_r)		
		LI	LS	Est.
4 a 8	Logístico	6,01	6,36	6,19
8 a 4	Gompertz	6,17	6,88	6,53

Para a ponta de pulverização de média vazão XR 110 04 tem-se resultados

semelhantes aos da ponta de baixa vazão. Observa-se da Tabela 4 que dentro de cada mudança de velocidade, a estimativa do tempo de resposta para o modelo de Gompertz é sempre maior que as estimativas correspondentes do modelo Logístico. Isso deve-se a fato do modelo de Gompertz ter uma curva assimétrica à direita, o que faz torná-lo mais “lento” para atingir a fração desejada da assíntota, e consequentemente, o tempo de resposta.

Tabela 4 – Estimativa (Est.), limite inferior (LI) e limite superior (LS) para o intervalo de confiança assintótico (95%) para o parâmetro de interesse tempo de resposta, t_r , dos modelos Logístico e de Gompertz para a ponta de média vazão

Velocidade (km/h)	Modelo ajustado	Tempo de Resposta (t_r)		
		LI	LS	Est.
12 a 16	Logístico	2,70	2,86	2,78
	Gompertz	2,82	3,04	2,93
16 a 12	Logístico	2,79	3,04	2,91
	Gompertz	3,05	3,35	3,20
12 a 18	Logístico	3,50	3,66	3,58
	Gompertz	3,75	3,99	3,87
18 a 12	Logístico	3,84	4,13	3,99
	Gompertz	4,33	4,60	4,47

Resultados semelhantes também são encontrados a respeito das medidas de não linearidade. Tanto CP e CI quanto o AIC indicam que o modelo melhor ajustado para as mudanças na ascendente é o logístico. Já na descendente, o modelo de Gompertz tem melhores resultados em relação aos critérios adotados (Tabela 5).

Tabela 5 – Medidas do critério de informação de Akaike (AIC) e de curvatura devido ao efeito de parametrização (CP) e intrínseca (CI) dos modelos logístico e de Gompertz para a ponta de média vazão

Velocidade (km/h)	Modelo	Curvatura		AIC
		CP	CI	
12 a 16	Logístico	0,0510	0,0532	459,35
	Gompertz	0,0526	0,0716	478,92
16 a 12	Logístico	0,0654	0,0555	490,26
	Gompertz	0,0517	0,0635	479,81
12 a 18	Logístico	0,0383	0,0344	490,70
	Gompertz	0,0434	0,0509	523,45
18 a 12	Logístico	0,0549	0,0451	528,73
	Gompertz	0,0358	0,0403	478,24

Na Figura 8, está a representação dos valores preditos para os modelos selecionados, juntamente com a banda de confiança (95%) e com destaque para a estimativa

intervalar assintótica (95%) do tempo de resposta de cada mudança realizada para a ponta de média vazão. Tem-se, que as diferenças entre os valores observados e preditos são menores na região próxima do ponto de estabilização para os dados na ascendente do que na descendente. Isso resulta em intervalos de confiança mais precisos para o tempo de resposta do sistema nas mudanças em que ocorre o aumento da velocidade.

Figura 8 – Valores observados e preditos com banda de confiança assintótica (95%) para a pressão (kPa) em função do tempo (s) para cada mudança de vazão. Linhas verticais identificam a estimativa intervalar assintótica (95%) do tempo de resposta do controlador eletrônico de pulverização para a ponta de média vazão XR 110 04

As estimativas pontuais e intervalares dos tempos de resposta para os modelos selecionados são apresentadas na Tabela 6. Observa-se que o tempo de resposta é menor para as ambas as mudanças feitas na ascendente. Da mesma tabela verifica-se que os tempos de resposta, para as mudanças de velocidade com amplitude de 6 km/h, são estatisticamente maiores que as mudanças com amplitude de 4 km/h, ou seja, o sistema

demora mais para estabilizar quando a amplitude da mudança de velocidade é maior utilizando-se a ponta de pulverização de média vazão.

Tabela 6 – Estimativa (Est.), limite inferior (LI) e limite superior (LS) para o intervalo de confiança assintótico (95%) para o parâmetro de interesse tempo de resposta, t_r , dos modelos selecionados para a ponta de média vazão

Velocidade (km/h)	Modelo	Tempo de Resposta (t_r)		
		LI	LS	Est.
12 a 16	Logístico	2,70	2,86	2,78
16 a 12	Gompertz	3,05	3,35	3,20
12 a 18	Logístico	3,50	3,66	3,58
18 a 12	Gompertz	4,33	4,60	4,47

Da Tabela 7 tem-se as estimativas do parâmetro de interesse para a ponta de pulverização de alta vazão XR 110 06. Verifica-se que os tempos de resposta estão muito próximos para ambos os modelos em ambas as mudanças em que a amplitude total é de apenas 0,43. Observa-se também que dentro de cada mudança os intervalos de confiança do parâmetro de interesse estão sobrepostos, ou seja, os tempos de resposta são estatisticamente iguais para os dois modelos propostos.

Tabela 7 – Estimativa (Est.), limite inferior (LI) e limite superior (LS) para o intervalo de confiança assintótico (95%) para o parâmetro de interesse tempo de resposta, t_r , dos modelos Logístico e de Gompertz para a ponta de alta vazão

Velocidade (km/h)	Modelo ajustado	Tempo de Resposta (t_r)		
		LI	LS	Est.
18 a 24	Logístico	2,61	2,79	2,70
	Gompertz	2,70	2,95	2,82
24 a 18	Logístico	2,78	3,02	2,90
	Gompertz	2,99	3,28	3,13

Igualmente às pontas anteriores, os modelos selecionados para as mudanças na ascendente e na descendente foram o logístico e Gompertz, respectivamente (Tabela 8). Observa-se que as medidas de não linearidade e o AIC estão muito próximas, mais mesmo assim tem-se uma ligeira vantagem à esses respectivos modelos.

Também para a ponta de alta vazão, a amplitude do intervalo de confiança para a mudança na descendente é maior que na ascendente, apesar dessa medida ser menor em comparação com as outras pontas de baixa e média vazão (Figura 9).

Tabela 8 – Medidas do critério de informação de Akaike (AIC) e de curvatura devido ao efeito de parametrização (CP) e intrínseca (CI) dos modelos logístico e de Gompertz para a ponta de alta vazão

Velocidade (km/h)	Modelo	Curvatura		AIC
		CP	CI	
18 a 24	Logístico	0,0591	0,0620	465,53
	Gompertz	0,0621	0,0853	488,00
24 a 18	Logístico	0,0610	0,0570	478,04
	Gompertz	0,0533	0,0692	477,74

Figura 9 – Valores observados e preditos com banda de confiança assintótica (95%) para a pressão (kPa) em função do tempo (s) para cada mudança de velocidade. Linhas verticais identificam a estimativa intervalar assintótica (95%) do tempo de resposta do controlador eletrônico de pulverização para a ponta de alta vazão XR 110 06

Apesar dos valores estimados para os tempos de resposta terem sido muito próximos para as mudanças, tem-se da Tabela 9 que foram estatisticamente diferentes. Assim, o sistema também é mais lento para corrigir a vazão quando há a redução da velocidade utilizando-se a ponta de alta vazão.

Tabela 9 – Estimativa (Est.), limite inferior (LI) e limite superior (LS) para o intervalo de confiança assintótico (95%) para o parâmetro de interesse tempo de resposta, t_r , dos modelos selecionados para a ponta de alta vazão

Velocidade (km/h)	Modelo	Tempo de Resposta (t_r)		
		LI	LS	Est.
18 a 24	Logístico	2,61	2,79	2,70
24 a 18	Gompertz	2,99	3,28	3,13

Por fim, os tempos de resposta obtidos no presente trabalho e os respectivos

encontrados em Vieira (2013), foram apresentados na Tabela 10. O modelo utilizado no trabalho anterior foi o logístico, em que as estimativas para o tempo de resposta foram obtidas por uma função do parâmetro x_I , ponto de inflexão. Assim, observa-se que a inclusão do parâmetro de interesse no modelo tem-se estimativas mais baixas para o tempo de resposta do sistema nos dados na ascendente.

Tabela 10 – Estimativa (Est.), limite inferior (LI), limite superior (LS) para o intervalo de confiança assintótico (95%) e estimativa de Vieira (2013) (Est. Antes), para o parâmetro de interesse tempo de resposta, t_r , para as três diferentes pontas

Ponta	Velocidade (km/h)	Tempo de Resposta (t_r)			
		LI	LS	Est.	Est. Antes
Baixa vazão	4 a 8	6,01	6,36	6,19	8,45
	8 a 4	6,17	6,88	6,53	5,53
Média vazão	12 a 16	2,70	2,86	2,78	3,25
	16 a 12	3,05	3,35	3,20	3,83
	12 a 18	3,50	3,66	3,58	4,42
	18 a 12	4,33	4,60	4,47	4,16
Alta vazão	18 a 24	2,61	2,79	2,70	3,56
	24 a 18	2,99	3,28	3,13	3,20

Agora, considerando que o modelo de Gompertz teve melhor ajuste para os dados na descendente, em que a curva não é simétrica em relação ao ponto de inflexão, x_I , tem-se que o tempo de resposta foi maior na ponta de baixa vazão em comparação com o trabalho anterior. Para ponta de média vazão, também foi menor na mudança de 18 a 12 km/h porém maior para a mudança de 16 a 12 km/h. Por fim, tem-se que o tempo de resposta foi estatisticamente igual para a ponta de alta vazão.

5.2 Avaliação dos dados simulados

Os resultados do estudo de simulação para avaliação dos estimadores dos parâmetros da equação (20) encontra-se na Tabela 11. Observa-se que as estimativas do viés dos parâmetros β_{20} , x_I , s , ν e τ são baixas, aproximando-se de zero à medida que o tamanho amostral é aumentado, como é esperado sob a teoria assintótica padrão. Em relação às estimativas dos EQMs, verifica-se que o único valor que não está abaixo de 1 é o do parâmetro x_I no tamanho amostral $n = 18$, em que o valor foi de 8,33, ficando bem acima em relação aos outros parâmetros. Mas, a medida que o tamanho amostral aumenta, todos os valores dos EQMs diminuem consideravelmente, aproximando-se de

zero, inclusive do parâmetro x_I . Esse fato indica consistência dos estimadores de máxima verossimilhança (EMV).

Ao considerar a propriedade de normalidade assintótica, espera-se que a distribuição dos estimadores se aproxime da distribuição normal, a medida que n aumenta. Baseado nessa propriedade, espera-se que as EMs dos parâmetros se aproximem dos verdadeiros valores e os coeficientes de assimetria e curtose se aproximem de zero e três, respectivamente. Ainda da Tabela 11, observa-se que as EMs dos parâmetros se aproximam dos verdadeiros valores conforme o tamanho amostral é aumentado, e mais, os valores dos coeficientes de assimetria (CA) e curtose (K) estão cada vez mais próximos dos valores zero e três, respectivamente.

Tabela 11 – EMs, Vies, EQMs, CA e K baseados em 1000 simulações do modelo de regressão não linear Exp-SC de parâmetros $\beta_{20} = 50$; $x_I = 50$; $s = 5$; $\nu = 0, 1$ e $\tau = 0, 6$

n	Parâmetro	EM	Viés	EQM	CA	K
18	β_{20}	50,015	0,015	0,743	0,210	3,103
	x_I	49,340	-0,660	8,330	-0,429	2,223
	s	4,693	-0,307	0,484	-0,391	3,059
	ν	0,080	-0,020	0,002	0,265	2,823
	τ	0,627	0,027	0,008	0,227	2,485
45	β_{20}	49,993	-0,007	0,287	0,158	3,386
	x_I	50,000	0,000	0,507	0,041	2,924
	s	4,982	-0,018	0,110	0,135	2,898
	ν	0,100	0,000	0,000	0,367	3,141
	τ	0,599	-0,001	0,001	0,050	2,978
135	β_{20}	50,019	0,019	0,101	0,134	3,152
	x_I	50,019	0,019	0,156	-0,037	3,312
	s	5,009	0,009	0,038	0,209	2,898
	ν	0,100	0,000	0,000	0,347	3,102
	τ	0,600	0,000	0,000	0,080	3,165
270	β_{20}	50,008	0,008	0,052	-0,061	3,082
	x_I	49,992	-0,008	0,077	0,176	3,278
	s	5,004	0,004	0,019	0,062	2,638
	ν	0,101	0,001	0,000	0,186	2,729
	τ	0,601	0,001	0,000	-0,081	3,111

Na Figura 10 tem-se os valores observados para os verdadeiros parâmetros selecionados e para as estimativas médias (EMs), em função da variável independente X . Observa-se que apenas no menor tamanho amostral, $n = 18$, que a curvas não se sobrepõem por completa.

Figura 10 – Valores observados para os verdadeiros parâmetros $\beta_{20} = 50$; $x_I = 50$; $s = 5$; $\nu = 0, 1$ e $\tau = 0, 6$ e para as estimativas médias (EMs), em função da variável independente x para os tamanhos amostrais: $n = 18; 45; 135; 270$

Figura 11 – Densidades estimadas de 1000 amostras para $n = 18; 45; 135; 270$ dos parâmetros $\beta_{20} = 50$; $x_I = 50$; $s = 5$; $\nu = 0, 1$ e $\tau = 0, 6$

As densidades estimadas com base em 1000 amostras das EMs dos parâmetros β_{20} , x_I , s , ν e τ foram representadas na Figura 11. Verifica-se que, para ambos os parâmetros, as propriedades inferenciais não foram muito precisas para o menor tamanho amostral, $n = 18$, o que já era esperado, devido à quantidade de parâmetros do modelo para apenas seis pontos amostrais, $h = 6$, sendo que a preocupação maior deve ser para o parâmetro x_I , indicando que a aproximação assintótica pode ser pobre nestes casos.

Mas, conforme o tamanho amostral aumenta, tem-se que as densidades

estimadas se aproximam cada vez mais da distribuição normal. Assim, tem-se que os gráficos apresentados estão de acordo com a teoria assintótica padrão para as estimativas de máxima verossimilhança.

6 CONSIDERAÇÕES FINAIS

Os modelos de regressão não linear logístico e de Gompertz são resultantes da soma ponderada de duas regressões lineares constantes com pesos dados pelas funções distribuição acumulada logística e Gumbel, respectivamente. Reparametrizações desses modelos foram feitas incluindo o tempo de resposta de um sistema de controle eletrônico de pulverização como parâmetro dos modelos. Para as mudanças na ascendente, o modelo selecionado, que atendeu as pressuposições, e pelas medidas de não linearidade e pelo critério de Akaike (AIC) foi o logístico. Já para as mudanças na descendente, o modelo selecionado foi o de Gompertz. Os tempos de resposta para a ponta de pulverização XR 110 02 (baixa vazão) são iguais para as duas mudanças de velocidade, sendo pouco mais de 6 segundos. Já para a ponta XR 110 04 (média vazão) houve diferença para as mudanças no aumento e na redução de velocidade para os dois saltos. Para o salto de 4 km/h a diferença pontual foi de 0,42 segundos. Já para o salto de 6 km/h foi de 0,89 segundos. Para a ponta XR 110 06 (alta vazão), essa diferença pontual foi menor, 0,43 segundos, mas também foi significativa. Os tempos de resposta diminuem conforme a ordem crescente da vazão nas pontas.

Por fim, considerando simulações de Monte Carlo, tem-se que os estimadores de máxima verossimilhança dos parâmetros do modelo de regressão não linear Cauchy seno hiperbólico exponencial possuem boas propriedades. Verificou-se que, mesmo em amostras pequenas, o viés relativo dos estimadores dos parâmetros são pequenos e que o EQM converge rapidamente à zero quando o tamanho amostral aumenta. Esse modelo é uma boa alternativa aos conhecidos na literatura como modelos de crescimento difásico, por ser mais flexível e ter um número menor de parâmetros.

6.1 Trabalhos futuros

Recomenda-se em trabalhos futuros a apresentação de um número maior de modelos de regressões não lineares resultantes da soma de regressões lineares ponderadas por funções distribuição acumulada que possivelmente resultam em modelos já conhecidos na literatura, e os que não resultam, que seja feito um estudo mais detalhado das interpretações dos parâmetros e justificativas para o uso da metodologia desenvolvida.

Aplicação do modelo de regressão não linear difásico Cauchy seno hiperbólico exponencial a um conjunto de dados reais e mostrar a aplicabilidade do mesmo.

REFERÊNCIAS

- AKAIKE, H. A. a new look at the statistical model identification. **IEEE Transation on Automatic Control**. New York, v. 19, p. 716-723, 1974.
- ALEXANDER, D.; RESTER, D.; MAYEUX, M.; THOMAS, C. Constant spray applications at varying speeds. **Lousiana Agriculture**. Lousiana, v. 7, n. 4, p. 8-9, 1964.
- ANTUNIASSI, U. R. **Avaliação de sistemas de injeção de defensivos para utilização em agricultura de precisão**, Botucatu, 1999. 83p. Tese livre Docência (Disciplina Mecanização Agrícola/ Departamento de Engenharia Rural) - Faculdade de Ciências Agrônômicas, Universidade Estadual Paulista “Júlio de Mesquita Filho”, 1999.
- BATES, D.M.; WATTS, D.G. Relative Curvature Measures of Nonlinearity. **Journal of the Royal Statistical Society**. Serie B. Methodological, London, v. 42, p. 1-25, 1980.
- BATES, D.M.; WATTS, D.G. **Nonlinear regression analysis and its applications**. New York: John Wiley & Sons. 1988. 365 p.
- BELSLEY, D.A; KUH, E. WELSCH, R.E. **Regression Diagnostics: Identifying Influential Data and Sources of Collinearity**. New York: John Wiley, 1980. 292 p.
- CARVALHO, W. P. A.; ANTUNIASSI, U. R. Avaliação de resposta de um controlador de fluxo com DGPS submetido à variação de velocidade de voo em condições de voo simulado em laboratório. **Energia na Agricultura**, Botucatu, vol. 27, n.2, abr./jun., p. 124-144, 2012.
- CLARKE, G. Approximate confidence limits for a parameter function in nonlinear regression. **Journal of the American Statistical Association**, New York, v. 82, n. 397, p. 221-230, Mar. 1987.
- COOK, R. D.; WEISBERG, S. **Residuals and influence in regression**. New York: Chapman and Hall, 1982. 240 p.
- COORAY, K. Exponentiated Sinh Cauchy Distribution with Applications. **Communications in Statistics-Theory and Methods**, Philadelphia, v. 42, p. 3838-3852, 2013.
- CUNHA, L.S. **Modelos de regressão para descrição do crescimento de frutos da ameixeira**. 2011. 80p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras, 2011.
- DRAPER, N.R; SMITH, H. **Applied regression analysis**. 3.ed. New York: J. Wiley & Sons, 1998. 706 p.
- FALCÃO, P. F. et al. Curvas de crescimento de cordeiros da raça Ile de France criados em confinamento. **Revista Brasileira de Saúde e Produção Animal**, Salvador, v. 16, n. 2, abr./jun., p. 377-386, 2015.
- FERNANDES, T. J. et al. Seleção de modelos não lineares para a descrição das curvas de crescimento do fruto do cafeeiro. **Coffee Science**, Lavras, v. 9, n. 2, abr./jun., p. 207-215, 2013.

FULTON, J.P.; SHEARER, S.A.; STOMBAUGH, T.S.; ANDERSON, M.E.; BURKS, T.F.; HIGGINS, S.F. Simulation of fixed-and variable-rate application of granular materials. **Transactions of ASAE**. St. Joseph, v. 46, n.5, p. 1311-1321, 2003.

FULTON, J. P.; SHEARER, S. A.; HIGGINS, S. F.; DARR, M. J.; STOMBAUGH, T. S. Rate response assessment from various granular VRT applicators. **Transactions of ASAE**. St. Joseph, v. 48, n.6, p. 2095-2103, 2005.

GADANHA JÚNIOR, C.D. **Avaliação do tempo de resposta de controladores eletrônicos em pulverizadores agrícolas**. Botucatu, 2000. 125p. Tese (Doutorado Energia na Agricultura) – Faculdade de Ciências Agronômicas, Universidade Estadual Paulista “Júlio de Mesquita Filho” 2000.

GALTON, F. **Natural Inheritance Macmillan**. London: Macmillan & Co, 1889. 259 p.

GOMPERTZ, B. On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies. **Philosophical Transactions of the Royal Society of London**. Londres, v. 115, p. 513-585, 1825.

GRAYBILL, F. A.; IYER, H. K. **Regression analysis: concepts and applications**. Belmont: Duxbury, 2004. 699 p.

GUPTA, R.D.; KUNDU, D. Exponentiated exponential family: an alternative to gamma and Weibull distributions. **Biometrical journal**, Malden, v. 43, p. 117-130, 2001.

MATTHEWS, D. Likelihood-based confidence intervals for functions of many parameters. **Biometrika**, London, v. 75, n. 1, p. 139-144, Sept. 1988.

MATUO, T; PIO, L.C.; RAMOS, H.H. **Módulo 2 - tecnologia de aplicação dos agroquímicos e equipamentos**. In: ASSOCIAÇÃO BRASILEIRA DE ENSINO SUPERIOR (ABEAS). Curso de especialização por tutoria à distância - Curso de proteção de plantas, Brasília: ABEAS, 2002. 91p.

MOHAMMADZAMANI, D. et al. Variable rate herbicide application using the global positioning system for generating a digital management map. **International Journal of Agriculture & Biology**. Pakistan, v. 11, n. 2, p. 178-182, 2009.

MORAIS, J. et al. Growth curve of different strains of broiler chickens. **Ciência Rural**, Santa Maria, v. 45, n. 10, p. 1872-1878, 2015.

MOTA, L. F. M. et al. Crescimento de codornas de diferentes grupos genéticos por meio de modelos não lineares. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, Belo Horizonte, v. 67, n. 5, p. 1372-1380, 2015.

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

PRAVEEN J.B. **Response time evaluation of real-time sensor based variable rate technology equipment**. 2009. p. 71. Dissertação (Mestrado Engenharia de biosistemas) - Oklahoma State University, Stillwater, 2009.

R Development Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL. Disponível em: <http://www.R-project.org/>. Acesso em: 09 jan. 2015.

RATKOWSKY, D.A. **Nonlinear regression modeling**: a unified practical approach. New York: Marcel Dekker, 1983. v. 48, 276 p.

RITZ, C.; STREIBIG, J. C. **Nonlinear regression with R**. New York: Springer, 2008. 142 p.

ROBINSON, A. P.; HAMANN, J. D. **Forest analytics with R**. New York: Springer, 2011. 354 p.

SCHABENBERGER, O.; PIERCE, F. J. **Contemporary statistical models for the plant and soil Sciences**. Boca Raton: CRC, 2002. 738 p.

SEBER, G. A. F.; WILD, C. J. **Nonlinear regression**. Hoboken: Wiley, 2003. 752 p.

SHARDA, A.; FULTON, J.P.; MCDONALD, T.P.; BRODBECK, C.J. Real-time nozzle flow uniformity when using automatic section control on agricultural sprayers. **Computers and Electronics in Agriculture**. Amsterdam, v. 79, p. 169-179, 2011.

SPRAYING SYSTEM. **TEEJET**. Teejet technologies: catalog 50A-P. Wheaton, 2011. p. 192.

VELOSO, R. C. et al. Crescimento de genótipos de frangos tipo caipira. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, Belo Horizonte, v. 67, n. 5, p. 1361-1371, 2015.

VENABLES, W. N.; RIPLEY, B. D. **Modern applied statistics with S**. 4ed. ed. New York: Springer, 2002. 495 p. .

VERHULST, P.F. Notice sur la loi que la population poursuit dans son accroissement. **Correspondence Mathématique et Physique**, Bruxelles, v. 10, p. 113-121, 1838.

VERION AGRICULTURA. **Vcom 8.4**: uso e manutenção, São Paulo, 2010. 20 p.

VIEIRA, R. R. **Tempo de resposta de um controlador eletrônico em sistemas de aplicação a taxas variáveis em pulverizações agrícolas**. 2013. 78 p. Dissertação (Mestrado em Máquinas Agrícolas) - Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 2013.

WINSOR, C.P. The Gompertz curve as a growth curve. **Proceedings of the National Academy of Sciences**. Washington, v. 18, p. 1-8, 1932.

ZEVIANI, W. M. **Parametrizações interpretáveis em modelos não linear**. 2013. 146p.
Tese (Tese de Doutorado) - Escola Superior de Agricultura Luiz de Queiroz, Universidade
Federal de Lavras, Lavras, 2013.

APÊNDICES

APÊNDICE A - Gráficos das análises dos resíduos para os modelos logístico e de Gompertz para cada ponta de pulverização

Figura 12 – Gráfico dos resíduos padronizados *vs* valores ajustados e o envelope simulado para a ponta de pulverização de XR 110 02 (baixa vazão) dos modelos logístico e de Gompertz

Figura 13 – Gráfico dos resíduos padronizados *vs* valores ajustados e o envelope simulado para a ponta de pulverização de XR 110 04 (média vazão) dos modelos logístico e de Gompertz

Figura 14 – Gráfico dos resíduos padronizados *vs* valores ajustados e o envelope simulado para a ponta de pulverização de XR 110 06 (alta vazão) dos modelos logístico e de Gompertz

No gráfico dos resíduos padronizados *versus* valores ajustados para cada ponta de pulverização (Figuras 12, 13 e 14), observa-se uma pequena tendência cíclica na dispersão dos pontos. Contudo, observa-se que os pontos estão distribuídos de forma mais aleatória para os modelos logísticos e de Gompertz nas mudanças realizadas na ascendente e descendente, respectivamente, e mais, eles estão dispersos no limite da faixa considerada ideal de -3 a 3, indicando que os ajustes destes modelos foram satisfatórios. Em relação às condições de normalidade dos erros, observa-se do gráfico meio-normal com envelope simulado (half-normal plot), que foram satisfeitas também para os modelos logísticos e de Gompertz para as mudanças na ascendente e descendente, respectivamente. As afirmações acerca disso também foram confirmadas pelo teste de Shapiro-Wilk, a um nível de 5% de significância (SHAPIRO e WILK, 1965).

APÊNDICE B - Parte do banco de dados para a ponta de pulverização XR 110 02 (baixa vazão)

Tabela 12 – Valores da pressão (Kpa) em função do tempo (s) da ponta de pulverização XR 110 02 (baixa vazão) para as mudanças realizadas

Ascendente			Descendente		
km/h	s	Kpa	km/h	s	Kpa
4	0,125	75,78383	8	0,125	168,8705
4	0,250	73,59793	8	0,250	170,1846
4	0,375	75,78383	8	0,375	166,6836
⋮	⋮	⋮	⋮	⋮	⋮
4	2,000	74,03530	8	2,000	165,3695
8	2,125	71,41301	4	2,125	170,6161
⋮	⋮	⋮	⋮	⋮	⋮
8	9,750	159,26000	4	9,750	74,9091
8	9,875	154,44493	4	9,875	74,9091
8	10,000	158,37740	4	10,000	76,2202

As tabelas do banco de dados para as pontas de pulverização de média e alta vazão são análogas à Tabela 12.

ANEXOS

ANEXO A - Códigos no software R para o ajuste e comparação de modelos não lineares logístico e de Gompertz para encontrar os tempos de resposta de um sistema de controle eletrônico de pulverização

```
##-----
### PONTA DE PULVERIZAÇÃO DE BAIXA VAZÃO XR 110 02 ###
## VELOCIDADE 4-8 km/h ##
rm(list=ls())
setwd("C:/Users/LucasPanda/Dropbox/Doutorado/tese.R/dados")
dados4_8 <- read.csv2("dados4_8.csv",h=T)
# Gráfico de dispersão dos dados #
plot(kPa~segundos,main=expression("4 a 8 km/h"),xlab="Tempo (s)",
      ylab="Pressão (kPa)",data=dados4_8)
##-----
## Modelos propostos ##
# Logístico reparametrizado ( isolando s)
logis_tr <- deriv3(~(beta1+
(beta2-beta1)/(1+(((1-q)*beta2)/(q*beta2-beta1))^(xI-x)/(xI-tm-tr)))),
      c("beta1","beta2","xI","tr"),
      function(x, beta1, beta2, xI, tr){ NULL })
# Gompertz reparametrizado (isolando s)
gomp_tr <- deriv3(~(beta1+
(beta2-beta1)*exp(-1/(-log((q*beta2-beta1)/(beta2-beta1)))^(x-xI)/(xI-tm-tr)))
      ),c("beta1","beta2","xI","tr"),
      function(x, beta1, beta2, xI, tr){ NULL })
##-----
## Ajuste dos modelos ##
q=0.95 #fixo
tm=2 #fixo
# Logístico reparametrizado
alogis_tr <- nls(kPa~logis_tr(segundos, beta1, beta2, xI, tr), data=dados4_8,
      start=c(beta1=80,beta2=160,xI=5,tr=6))
summary(alogis_tr)
confint.default(alogis_tr) # Intervalo de confiança assintótico
# gompertz reparametrizado
agomp_tr <- nls(kPa~gomp_tr(segundos, beta1, beta2, xI, tr), data=dados4_8,
      start=c(beta1=80,beta2=160,xI=5,tr=6))
summary(agomp_tr)
confint.default(agomp_tr) # Intervalo de confiança assintótico
##-----
## Medidas de curvatura ##
rms.curv(alogis_tr)
rms.curv(agomp_tr)
# AIC E BIC
AIC(alogis_tr);AIC(agomp_tr)
#-----
#-----
## Análise dos resíduos ##
# Modelo logístico
# extraíndo a matriz gradiente avaliada nas estimativas dos parâmetros
F <- attr(alogis_tr$m$fitted(), "gradient")
```

```

# passando a matriz gradiente para a lm(), importante remover intercepto
m0 <- lm(kPa~-1+F, data=dados4_8)
# gráfico de análise dos resíduos
# gráfico de análise dos resíduos
r <- rstandard(m0)
f <- fitted(m0)
par(mfrow=c(1,2), mar=c(5,4,4,2))
plot(r~f, ylim=c(-4,4),xlab="Valores ajustados", ylab="Resíduos padronizados")
abline(h=-3,lty=2)
abline(h=3,lty=2)
hnp(m0, xlab="Quantis teóricos",ylab="Resíduos")
mtext("Velocidade 4 a 8 km/h - Modelo logístico",
outer=TRUE, line=-2, cex=1)
layout(1)
# testes
bgtest(m0)
dwt(m0)
shapiro.test(r)
# Modelo de Gompertz
# extraindo a matriz gradiente avaliada nas estimativas dos parâmetros
F <- attr(agomp_tr$m$fitted(), "gradient")
# passando a matriz gradiente para a lm(), importante remover intercepto
m0 <- lm(kPa~-1+F, data=dados4_8)
# gráfico de análise dos resíduos
r <- rstandard(m0)
f <- fitted(m0)
par(mfrow=c(1,2), mar=c(5,4,4,2))
plot(r~f, ylim=c(-4,4),xlab="Valores ajustados", ylab="Resíduos padronizados")
abline(h=-3,lty=2)
abline(h=3,lty=2)
hnp(m0, xlab="Quantis teóricos",ylab="Resíduos")
mtext("Velocidade 4 a 8 km/h - Modelo de Gompertz",
outer=TRUE, line=-2, cex=1)
# testes
dwtest(m0)
shapiro.test(r)
##-----
# Obtenção dos valores preditos
pred <- data.frame(segundos=seq(0.125,10,1=1000))
der <- do.call(logis_tr, args=c(list(x=pred$segundos),as.list(coef(alogis_tr))))
F <- attr(der, "gradient") # gradiente avaliado no novo t
U <- chol(vcov(alogis_tr))
se <- sqrt(apply(F%*%t(U), 1, function(x) sum(x^2))) # erro padrão
#-----
# gráficos dos observados, preditos com IC
plot(kPa~segundos, data=dados4_8, col=1, main=expression("4 a 8 km/h"),
      xlab="Tempo (s)", ylab="Pressão (kPa)")
matlines(pred$segundos, c(der)+
          outer(se, qt(c(.5, .025,.975), df=df.residual(alogis_tr))),
          type="l", col=c(3,2,2), lty=c(1,2,2))
layout(1)

```

```

abline(v=2+6.19)
abline(v=2+6.01,lty=2)
abline(v=2+6.36,lty=2)
##-----
##-----
### VELOCIDADE 8-4 km/h ###
rm(list=ls())
setwd("C:/Users/LucasPanda/Dropbox/Doutorado/tese.R/dados")
dados4_8 <- read.csv2("dados8_4.csv",h=T)
# Gráfico de dispersão dos dados #
plot(kPa~segundos,main=expression("8 a 4 km/h"),xlab="Tempo (s)",
      ylab="Pressão (kPa)",data=dados8_4)
##-----
## Modelos propostos ##
# Logístico reparametrizado ( isolando s)
logis_tr <- deriv3(~(beta1+
(beta2-beta1)/(1+(((1-q)*beta2)/(q*beta2-beta1))^(xI-x)/(xI-tm-tr)))),
      c("beta1","beta2","xI","tr"),
      function(x, beta1, beta2, xI, tr){ NULL })
# Gompertz reparametrizado (isolando s)
gomp_tr <- deriv3(~(beta1+
(beta2-beta1)*exp(-1/(-log((q*beta2-beta1)/(beta2-beta1)))^(x-xI)/(xI-tm-tr)))
      ),c("beta1","beta2","xI","tr"),
      function(x, beta1, beta2, xI, tr){ NULL })
##-----
## Ajuste dos modelos ##
q=1.05 #fixo
tm=2 #fixo
# Logístico reparametrizado
alogis_tr <- nls(kPa~logis_tr(segundos, beta1, beta2, xI, tr), data=dados8_4,
      start=c(beta1=160,beta2=80,xI=5,tr=6))
summary(alogis_tr)
confint.default(alogis_tr) # Intervalo de confiança assintótico
# gompertz reparametrizado
agomp_tr <- nls(kPa~gomp_tr(segundos, beta1, beta2, xI, tr), data=dados8_4,
      start=c(beta1=160,beta2=80,xI=5,tr=6))
summary(agomp_tr)
confint.default(agomp_tr) # Intervalo de confiança assintótico
##-----
## Medidas de curvatura ##
rms.curv(alogis_tr)
rms.curv(agomp_tr)
# AIC E BIC
AIC(alogis_tr);AIC(agomp_tr)
##-----
## Análise dos resíduos ##
# Modelo logístico #
# extraíndo a matriz gradiente avaliada nas estimativas dos parâmetros
F <- attr(alogis_tr$m$fitted(), "gradient")
# passando a matriz gradiente para a lm(), importante remover intercepto
m0 <- lm(kPa~-1+F, data=dados8_4)

```



```

# gráfico de análise dos resíduos
r <- rstandard(m0)
f <- fitted(m0)
par(mfrow=c(1,2), mar=c(5,4,4,2))
plot(r~f, ylim=c(-4,4),xlab="Valores ajustados", ylab="Resíduos padronizados")
abline(h=-3,lty=2)
abline(h=3,lty=2)
hnp(m0, xlab="Quantis teóricos",ylab="Resíduos")
mtext("Velocidade 8 a 4 km/h - Modelo logístico",
outer=TRUE, line=-2, cex=1)
layout(1)
# testes
dwt(m0)
shapiro.test(r)
# Modelo de Gompertz #
# extraíndo a matriz gradiente avaliada nas estimativas dos parâmetros
F <- attr(agomp_tr$m$fitted(), "gradient")
# passando a matriz gradiente para a lm(), importante remover intercepto
m0 <- lm(kPa~-1+F, data=dados8_4)
# gráfico de análise dos resíduos
r <- rstandard(m0)
f <- fitted(m0)
par(mfrow=c(1,2), mar=c(5,4,4,2))
plot(r~f, ylim=c(-4,4),xlab="Valores ajustados", ylab="Resíduos padronizados")
abline(h=-3,lty=2)
abline(h=3,lty=2)
hnp(m0, xlab="Quantis teóricos",ylab="Resíduos")
mtext("Velocidade 8 a 4 km/h - Modelo de Gompertz",
outer=TRUE, line=-2, cex=1)
# testes
dwt(m0)
shapiro.test(r)
##-----
# Obtenção dos valores preditos
pred <- data.frame(segundos=seq(0.125,10,1=1000))
der <- do.call(gomp_tr, args=c(list(x=pred$segundos), as.list(coef(agomp_tr))))
F <- attr(der, "gradient") # gradiente avaliado no novo t
U <- chol(vcov(agomp_tr))
se <- sqrt(apply(F%*%t(U), 1, function(x) sum(x^2))) # erro padrão
#-----
# gráficos dos observados, preditos com IC
plot(kPa~segundos, data=dados8_4, col=1, main=expression("8 a 4 km/h"),
      xlab="Tempo (s)", ylab="Pressão (kPa)")
matlines(pred$segundos, c(der)+
          outer(se, qt(c(.5, .025,.975), df=df.residual(agomp_tr))),
          type="l", col=c(3,2,2), lty=c(1,2,2))
layout(1)
abline(v=2+6.53)
abline(v=2+6.17,lty=2)
abline(v=2+6.88,lty=2)
##-----

```

###-----
OS CÓDIGOS PARA AS PONTAS DE PULVERIZAÇÃO DE MÉDIA E ALTA VAZÃO SÃO DADOS DE
FORMA ANÁLOGA AOS DESCRITOS ANTERIORMENTE.
###-----

ANEXO B - Códigos no software R para o estudo de simulação do modelo não linear difásico Cauchy seno hiperbólico exponencial, para mostrar a consistência e validade do mesmo

```
###-----
## SIMULAÇÃO ####
### -----
rm(list=ls(all=TRUE))
require(lattice)
require(car)
require(bbmle)
require(moments)
# Modelo de regressão não linear Cauchy seno hiperbólico exponencial (NLEXP-SC)
fac_noia <- function(x,beta20,xI,s,v,tau){
  (a)*(.5+(1/pi)*atan(v*sinh((x-xI)/s)))^tau
}
# Função de Verossimilhança.
llmod <- function(beta20,xI,s,v,tau,x,y){
  ex <- beta20*(.5+(1/pi)*atan(v*sinh((x-xI)/s)))^tau
  sd <- sqrt(crossprod(y-ex)/(length(ex)))
  ll <- sum(dnorm(y, mean=ex, sd=sd, log=TRUE))
  -ll
}
# Função que realiza a simulação
eqm<-function(h,r,beta201,xI1,s1,v1,tau1,casa.dec){
  beta20<-c();xI<-c();s<-c();v<-c();tau<-c()
  plot(0,0,ylim=c(0,1000),xlim=c(0,1))
  for (i in 1:1000){
    abline(h=i)
    x <- rep(seq(7,90,length.out = h),e=r)
    y <- rnorm(length(x),fac_noia(x,beta201,xI1,s1,v1,tau1),1.5)
    dados <- data.frame(x,y)
    alnoia <- mle2(minuslogl=llmod,
                  start=list(beta20=beta201,xI=xI1,s=s1,v=v1,tau=tau1),
                  data=dados, method="BFGS")
    beta20<-c(beta20,coef(alnoia)[1])
    xI<-c(xI,coef(alnoia)[2])
    s<-c(s,coef(alnoia)[3])
    v<-c(v,coef(alnoia)[4])
    tau<-c(tau,coef(alnoia)[5])
  }
  means<-c(mean(beta20),mean(xI),mean(s),mean(v),mean(tau))
  vars<-c(var(beta20),var(xI),var(s),var(v),var(tau))
  vies<-means-c(beta201,xI1,s1,v1,tau1)
  eqms<-vars+vies^2
  assim <- c(skewness(beta20),skewness(xI),skewness(s),skewness(v),
            skewness(tau))
  curtose <- c(kurtosis(beta20),kurtosis(xI),kurtosis(s),kurtosis(v),
            kurtosis(tau))
  result<-new.env()
  result$means<-round(means,casa.dec)
```

```

result$vars<-round(vars,casa.dec)
result$vies<-round(vies,casa.dec)
result$eqms<-round(eqms,casa.dec)
result$assim<-round(assim,casa.dec)
result$curtose<-round(curtose,casa.dec)
result$beta20<-beta20
result$xI<-xI
result$s<-s
result$v<-v
result$tau<-tau
t<-seq(7,90,0.01)
p1<-plot(x,y,type="p",ylab=expression(" "*eta(x)),xlab="x",cex=0.5)
p2<-lines(t,fac_noia(t,beta201,xI1,s1,v1,tau1),lwd=3)
p3<-lines(t,fac_noia(t,mean(beta20),mean(xI),mean(s),mean(v),mean(tau)),
          lwd=3,lty=2,col='brown1')
p4<-legend('right',c("Verdadeiro","Médio"),col=c(1,'brown1'),bty="n",
          lwd=2,lty=c(1,2),seg.len =1.5,cex=1.5)
p5<-legend('topleft',c(paste("h=",h),paste("r=",r)),col=c(0),
          bty="n",cex=1.5)
return(as.list(result))
return(p1)
return(p2)
return(p3)
return(p4)
return(p5)
}
###-----
# Simulação para h=6, r=3, então n=18.
setEPS()
postscript("C:/Users/LucasPanda/Dropbox/Doutorado/Tese Lucas/simul_18.eps")
set.seed(1)
test<-eqm(6,3, 50,50,5,0.1,.6, 3)
dev.off()
test$means
test$vies
test$eqms
test$assim
test$curtose
# Simulação para h=15, r=3, então n=45.
setEPS()
postscript("C:/Users/LucasPanda/Dropbox/Doutorado/Tese Lucas/simul_45.eps")
set.seed(2)
test2<-eqm(15,3, 50,50,5,0.1,.6, 3)
dev.off()
test2$means
test2$vies
test2$eqms
test2$assim
test2$curtose
# Simulação para h=45, r=3, então n=135.
setEPS()

```

```

postscript("C:/Users/LucasPanda/Dropbox/Doutorado/Tese Lucas/simul_135.eps")
set.seed(3)
test3<-eqm(45,3, 50,50,5,0.1,.6, 3)
dev.off()
test3$means
test3$vies
test3$eqms
test3$assim
test3$curtose
# Simulação para h=90, r=3, então n=270.
setEPS()
postscript("C:/Users/LucasPanda/Dropbox/Doutorado/Tese Lucas/simul_270.eps")
set.seed(270)
test4<-eqm(90,3, 50,50,5,0.1,.6, 3)
dev.off()
test4$means
test4$vies
test4$eqms
test4$assim
test4$curtose
###-----
# Gráfico das densidades estimadas para n=18; 45; 135; 270
de cada parâmetro ajustado
# Para o parâmetro beta20 #
setEPS()
postscript("C:/Users/LucasPanda/Dropbox/Doutorado/Tese Lucas/histbeta0.eps")
plot(density(test$beta20,bw=.3),lwd=3,col='chocolate1',
      xlim=c(48,52),ylim=c(0,1.3),xlab="",ylab="",main="")
lines(density(test2$beta20,bw=.3),col='chartreuse2',lwd=3)
lines(density(test3$beta20,bw=.3),col='dodgerblue1',lwd=3)
lines(density(test4$beta20,bw=.3),col=6,lwd=3)
abline(v=50,lwd=2,lty=2,col=1)
legend("topright",c('n=18','n=45','n=135','n=270',"True value"),
      col=c('chocolate1','chartreuse2','dodgerblue1',6,1),
      bty="n",seg.len=2,cex=1.5,lwd=3,lty=c(1,1,1,1,2))
mtext(expression(beta[20]), side=1, line=3, col="1", cex=2)
mtext("Densidade", side=2, line=3, col="1", cex=2)
dev.off()
# Para o parâmetro xI #
setEPS()
postscript("C:/Users/LucasPanda/Dropbox/Doutorado/Tese Lucas/histmu.eps")
plot(density(test$xI),ylim=c(0,.8),xlim=c(46,54),col='chocolate1',
      lwd=3,xlab="",ylab="",main="")
lines(density(test2$xI,bw=.5),col='chartreuse2',lwd=3)
lines(density(test3$xI,bw=.5),col='dodgerblue1',lwd=3)
lines(density(test4$xI,bw=.5),col=6,lwd=3)
abline(v=50,lwd=2,lty=2,col=1)
legend("topright",c('n=18','n=45','n=135','n=270',"True value"),
      col=c('chocolate1','chartreuse2','dodgerblue1',6,1),
      bty="n",seg.len=2,cex=1.5,lwd=3,lty=c(1,1,1,1,2))
mtext(expression(x[I]), side=1, line=3, col="1", cex=1.8)

```

```

mtext("Densidade", side=2, line=3, col="1", cex=1.6)
dev.off()
# Para o parâmetro s #
setEPS()
postscript("C:/Users/LucasPanda/Dropbox/Doutorado/Tese Lucas/histsigma.eps")
plot(density(test$s,bw=.5),ylim=c(0,.8),col='chocolate1',
      lwd=3,xlab="",main="",ylab="")
lines(density(test2$s,bw=.5),col='chartreuse2',lwd=3)
lines(density(test3$s,bw=.5),col='dodgerblue1',lwd=3)
lines(density(test4$s,bw=.5),col=6,lwd=3)
abline(v=5,lwd=2,lty=2,col=1)
legend("topright",c('n=18','n=45','n=135','n=270',"True value"),
      col=c('chocolate1','chartreuse2','dodgerblue1',6,1),
      bty="n",seg.len=2,cex=1.5,lwd=3,lty=c(1,1,1,1,2))
mtext(expression(s), side=1, line=3, col="1", cex=1.8)
mtext("Densidade", side=2, line=3, col="1", cex=1.6)
dev.off()
# Para o parâmetro v #
setEPS()
postscript("C:/Users/LucasPanda/Dropbox/Doutorado/Tese Lucas/histnu.eps")
plot(density(test$v,bw=.03),lwd=3,col='chocolate1',
      xlab="",ylab="",ylim=c(0,15),main="")
lines(density(test2$v,bw=.03),col='chartreuse2',lwd=3)
lines(density(test3$v,bw=.03),col='dodgerblue1',lwd=3)
lines(density(test4$v,bw=.03),col=6,lwd=3)
abline(v=0.1,lwd=2,lty=2,col=1)
legend("topright",c('n=18','n=45','n=135','n=270',"True value"),
      col=c('chocolate1','chartreuse2','dodgerblue1',6,1),
      bty="n",seg.len=2,cex=1.5,lwd=3,lty=c(1,1,1,1,2))
mtext(expression(nu), side=1, line=3, col="1", cex=1.8)
mtext("Densidade", side=2, line=3, col="1", cex=1.6)
dev.off()
# Para o parâmetro tau #
setEPS()
postscript("C:/Users/LucasPanda/Dropbox/Doutorado/Tese Lucas/histtau.eps")
plot(density(test$tau,bw=.06),ylim=c(0,7),col='chocolate1',lwd=3,
      xlab="",main="",ylab="")
lines(density(test2$tau,bw=.06),col='chartreuse2',lwd=3)
lines(density(test3$tau,bw=.06),col='dodgerblue1',lwd=3)
lines(density(test4$tau,bw=.06),col=6,lwd=3)
abline(v=.6,lwd=2,lty=2,col=1)
legend("topright",c('n=18','n=45','n=135','n=270',"True value"),
      col=c('chocolate1','chartreuse2','dodgerblue1',6,1),
      bty="n",seg.len=2,cex=1.5,lwd=3,lty=c(1,1,1,1,2))
mtext(expression(tau), side=1, line=3, col="1", cex=1.8)
mtext("Densidade", side=2, line=3, col="1", cex=1.6)
dev.off()
###-----

```