

**MODELOS MISTOS LINEARES EM ANÁLISE DE DADOS
LONGITUDINAIS EM GENÓTIPOS DE FEIJÃO VAGEM**

WANDERSON SOUZA RABELLO

**UNIVERSIDADE ESTADUAL DO NORTE FLUMINENSE
DARCY RIBEIRO - UENF**

**CAMPOS DOS GOYTACAZES – RJ
JUNHO – 2015**

**MODELOS MISTOS LINEARES EM ANÁLISE DE DADOS
LONGITUDINAIS EM GENÓTIPOS DE FEIJÃO VAGEM**

WANDERSON SOUZA RABELLO

Tese apresentada ao Centro de Ciências e Tecnologias Agropecuárias da Universidade Estadual do Norte Fluminense Darcy Ribeiro, como parte das exigências para obtenção do título de Doutor em Produção Vegetal.

Orientador: José Tarcísio Lima Thiebaut

**CAMPOS DOS GOYTACAZES - RJ
JUNHO – 2015**

FICHA CATALOGRÁFICA

Preparada pela Biblioteca do **CCTA / UENF** 181/2015

Rabello, Wanderson Souza

Modelos mistos lineares em análise de dados longitudinais em genótipos de feijão vagem / Wanderson Souza Rabello. – 2015.
108 f.

Orientador: José Tarcísio Lima Thiebaut

Tese (Doutorado - Produção Vegetal) – Universidade Estadual do Norte Fluminense Darcy Ribeiro, Centro de Ciências e Tecnologias Agropecuárias. Campos dos Goytacazes, RJ, 2015.

Bibliografia: f. 90 – 108.

1. Análise univariada 2. Análise multivariada 3. Estruturas de covariâncias 4. Dados correlacionados 5. Feijão-de-vagem I. Universidade Estadual do Norte Fluminense Darcy Ribeiro. Centro de Ciências e Tecnologias Agropecuárias. II. Título.

CDD– 635.651

MODELOS MISTOS LINEARES EM ANÁLISE DE DADOS
LONGITUDINAIS EM GENÓTIPOS DE FEIJÃO VAGEM

WANDERSON SOUZA RABELLO

Tese apresentada ao Centro de Ciências e Tecnologias Agropecuárias da Universidade Estadual do Norte Fluminense Darcy Ribeiro, como parte das exigências para obtenção do título de Doutor em Produção Vegetal.

Aprovada em 30 de junho de 2015.

Comissão examinadora

Prof. Rogério Figueiredo Daher (D. Sc., em Produção Vegetal) - UENF

Prof. Geraldo de Amaral Gravina (D. Sc., em Fitotecnia) - UENF

Leonardo Siqueira Glória (D. Sc., em Genética e Melhoramento) - UFV

Prof. José Tarcísio Lima Thiebaut (D. Sc. Em Ciência Animal) – UENF
(Orientador)

*Penso que cumprir a vida
Seja simplesmente
Compreender a marcha
E ir tocando em frente*

*Cada um de nós compõe a sua história
Cada ser em si
Carrega o dom de ser capaz
E ser feliz*

“Almir Sater”

AGRADECIMENTOS

A Deus que é merecedor de todo o meu louvor;

A Universidade Estadual do Norte Fluminense “Darcy Ribeiro” – UENF, por possibilitar a realização desse trabalho e pela concessão da bolsa de estudos;

Ao Prof. José Tarcísio Lima Thiebaut pela amizade, pelos ensinamentos, pela paciência, pela orientação e pelo apoio durante todo o curso;

Ao José Francisco Sá Vasconcelos Júnior pela amizade, pela força e pelo companheirismo;

A Igreja Cristã Maranata de Alphaville – Campos dos Goytacazes, pelo acolhimento, pelas alegrias, pela força, pelas experiências alcançadas nesses sete anos de convívio;

Aos meus pais, Antônio e Maria da Penha Rabello e a toda a família Rabello pelo apoio, pelas orações e pela ajuda;

Ao amor da minha vida, Danuza Barbosa, pela paciência, pelas orações e por ter estado ao meu lado sempre;

Aos amigos e colegas da Pós-Graduação em Produção Vegetal.

SUMÁRIO

RESUMO	vi
ABSTRACT	vii
1.0. INTRODUÇÃO	1
2.0. REVISÃO DE LITERATURA	3
2.1. Feijão vagem	3
2.1.1. Origem	3
2.1. 2. Botânica e características	4
2.1. 3. Importância econômica da cultura	6
2.1.4. Aspectos sobre a produção de feijão vagem	7
2.1.4.1. Clima	
2.1.4.2. Época de plantio e cultivo protegido	9
2.1.4.3. Solo e adubação	11
2.2.1. Dados longitudinais	13
2.2.2. Análise univariada de perfil	18
2.2.2.1 Correção dos graus de liberdade	23
2.2.3. Análise multivariada de perfil	25
2.2.4. Modelos Mistos	27
2.2.4.1. Especificação do modelo misto	29
2.2.4.2. Estimacão de β e Predicão de b	31
2.2.4.3. Estimacão dos parâmetros de covariância em G e R	34

2.2.4.4. Estruturas de Covariância	36
2.2.5. Seleção de Modelos	41
2.2.6. Teste da Razão de Verossimilhança	42
3.0. MATERIAL E MÉTODOS	44
3.1. Programa de melhoramento do feijão vagem da Universidade Estadual do Norte Fluminense Darcy Ribeiro (UENF)	
3.2. Material	45
3.3. Métodos	46
3.3.1. Análise Univariada de Perfil	47
3.2.1. Análise multivariada de perfil	48
3.3.1. Análise via modelos mistos	49
4.0 RESULTADOS E DISCUSSÃO	52
4.1 Análise Exploratória	52
4.2 Análise univariada de perfil	53
4.3 Análise multivariada de perfil	57
4.4. Modelo Misto	60
4.4.1. Modelo Misto com efeitos aleatórios de intercepto, tempo e tempo ² (U)	60
4.4.2. Modelo Misto com efeitos aleatórios de tempo e tempo ² (E)	63
4.4.3. Modelo Misto com efeito aleatório de tempo ² (N)	65
4.4.4. Modelo Misto sem efeitos aleatórios (F)	68
4.4.5. Comparação dos modelos selecionados dentro de cada grupo (UxExNx \bar{F}) formado	70
5. RESUMOS E CONCLUSÕES	75
6.0 REFERÊNCIAS BIBLIOGRÁFICAS	77

RESUMO

Rabello, Wanderson Souza. Universidade Estadual do Norte Fluminense Darcy Ribeiro. Junho de 2015. Modelos mistos lineares em análise de dados longitudinais em genótipos de feijão vagem. Orientador: Prof. José Tarcísio Lima Thiebaut.

O objetivo da pesquisa foi o estudo de dados longitudinais de genótipos de feijão vagem por meio da análise univariada de perfil, multivariada de perfil e modelos mistos. Os dados utilizados foram provenientes de um experimento realizado no Instituto Federal Fluminense, campus de Bom Jesus do Itabopoana – RJ, Campos dos Goytacazes – RJ. O experimento consistiu na avaliação de 17 genótipos de feijão vagem (*Phaseolus vulgaris* L.) de hábito de crescimento indeterminado do programa de melhoramento de feijão vagem da Universidade Estadual do Norte Fluminense Darcy Ribeiro - UENF, Campos dos Goytacazes – RJ. O trabalho foi implantado em um delineamento de blocos ao acaso com quatro repetições e 10 plantas por parcela no espaçamento de 1,0 x 0,5 m. A variável utilizada foi a Altura de Plantas, mensurada em centímetros (cm) em intervalos de quatro dias com régua de madeira de dois metros até os 36 dias após a semeadura. As avaliações realizadas caracterizaram as medidas no tempo, sendo, portanto 9 medidas no tempo. No presente trabalho verificou-se que a análise univariada de perfil não deve ser utilizada, pois não atendeu a condição H-F. Além disso, as correções dos graus de liberdade não promoveram

alterações nos resultados obtidos com a análise univariada de perfil. A análise multivariada mostrou-se uma alternativa correta na análise dos dados longitudinais por levar em consideração uma estrutura de covariância que considera a correlação entre os dados, porém a utilização dos modelos mistos se apresenta como uma alternativa mais versátil por poder entre outras vantagens testar diversas estruturas de covariâncias. Considerando o modelo misto, a inclusão da variação entre – indivíduos, ou seja, dos efeitos de intercepto, tempo e tempo² não melhorou a qualidade de ajuste. O melhor modelo foi ajustado com todos os efeitos fixos propostos, porém sem a variação entre – indivíduos e utilizando-se de uma estrutura de covariância não estruturada.

ABSTRACT

Rabello, Wanderson Souza. Universidade Estadual do Norte Fluminense Darcy Ribeiro. June, 2010. Linear mixed models for analyzing longitudinal data in snap bean genotypes. Advisor: Prof. José Tarcísio Lima Thiebaut.

The objective of the research was the study of longitudinal data snap bean genotypes pod through the profile univariate analysis, multivariate profile and mixed models. The data used came from an experiment performed at the North part of Rio de Janeiro State Federal Institute campus of Bom Jesus do Itabopoana - RJ, Campos dos Goytacazes - RJ. The experiment consisted of the evaluation of 17 pod snap bean genotypes (*Phaseolus vulgaris* L.) of indeterminate growth habit of the North part of Rio de Janeiro State University Darcy Ribeiro pod snap bean breeding program - UENF, Campos dos Goytacazes - RJ. The work was implemented on a randomized block design with four replications and 10 plants per plot in the spacing of 1.0 x 0.5 m. The variable used was the plant height, measured in centimeters (cm) at intervals of four days with wooden ruler from two meters to 36 days after sowing. The evaluations performed characterized the measures in time, and therefore 9 measures in time. In this study it was found that the univariate analysis profile should not be used because it did not meet the HF condition. In addition, the correction of degrees of freedom did not promote changes in the results of the univariate analysis profile. Multivariate analysis

proved to be a correct alternative in the analysis of longitudinal data by taking into account a covariance structure that considers the correlation between the data, but the use of mixed models is presented as a more versatile alternative for power among other advantages test different covariance structures. Considering the mixed model, the inclusion of variation between - individuals, that is, the effects of intercept time and tempo² not improve the quality of fit. The model was fit best with all the proposed purposes fixed, however without variation between - individuals and using an unstructured covariance structure.

1.0. INTRODUÇÃO

O feijão de vagem (*Phaseolus vulgaris* L.), também conhecido como vagem, é uma hortaliça de interesse mundial, pois tem relevante importância na nutrição humana como fonte de proteína, vitamina A, B1, B2 e C, sais minerais como cálcio, fósforo, potássio, ferro e por apresentar elevado teor de fibras (Hervantin e Teixeira, 1999; Figueira, 2000; Peixoto et al., 2002). Apesar da existência de cultivares específicos, qualquer cultivar de feijão pode ser usado para colheita das vagens (Vieira, 2006), que devem ser colhidas em seu ponto máximo de desenvolvimento, antes que se tornem fibrosas e com sementes salientes (Abreu et al., 2004).

Embora exista no mercado brasileiro cultivares de boa aceitação comercial, não há um programa nacional de avaliação e recomendação de genótipos mais produtivos em cada ambiente específico (Hamasaki et al., 1998). Pesquisas visando o melhoramento do feijão vagem são de elevada importância para que se possam obter melhorias no manejo e na produtividade da cultura (Kurek et al., 2001). Nesse sentido a Universidade Estadual do Norte Fluminense Darcy Ribeiro (UENF) por meio do programa de melhoramento do feijão vagem de hábito indeterminado, tem desenvolvido pesquisas com a finalidade de desenvolver cultivares de feijão vagem, visando contribuir com o aumento de produtividade, estabilidade e qualidade, além de reduzir os impactos ambientais e os custos de produção (Almeida, 2011).

Os estudos envolvendo dados longitudinais têm como objetivos o estudo do comportamento da variável resposta ao longo do tempo e a verificação da existência de dependência da variável resposta em relação às covariáveis (Wald,

2000; Costa, 2003; Amador, 2010). Resende et al. (1999) complementam que a interação época x tratamento é, provavelmente, o mais importante aspecto destes estudos, e quando significativa, indica uma tendência não paralela da resposta em relação ao tempo para diferentes níveis de tratamentos.

A característica distinta de estudos longitudinais é a dimensão ordenada com que os dados são coletados e o fato de que as observações repetidas para um indivíduo tendem a ser correlacionadas. Tal correlação pode ser modelada através de uma estrutura de covariância dos dados observados sendo que, para outros tipos de dados, é usual assumir que os erros sejam independentes. Modelar uma estrutura de covariâncias apropriada é essencial para que as inferências sobre as médias sejam válidas (Costa, 2003).

Para avaliar o efeito da correlação entre as observações nos diferentes tempos, podem-se analisar os dados longitudinais pelos seguintes modelos: univariado, multivariado e misto. Freitas (2007) afirma que o modelo univariado, conforme o planejamento de parcelas subdivididas no tempo impõe forte restrição à matriz de variâncias e covariâncias dos dados; o modelo multivariado utiliza somente uma matriz de covariância não estruturada; e, por último, o modelo misto permite a seleção de uma matriz com melhor representatividade dos dados.

Segundo Amador (2010), uma alternativa para a análise desse tipo de dado são as curvas de crescimento por meio de modelos mistos, pois leva em consideração a correlação serial dos dados.

A análise de crescimento é um método de grande utilidade para a avaliação das diferenças no comportamento de cultivares influenciadas por práticas agrônômicas, efeitos de competição ou climáticos, e por fatores intrínsecos associados à fisiologia da planta (Andrade et al., 2005; Guimarães et al., 2008). Os princípios e as práticas da análise de crescimento têm como objetivo descrever e interpretar o desempenho de determinada espécie em ambiente natural ou controlado (Hunt, 1990).

O presente trabalho teve como objetivo o emprego da análise univariada, multivariada e o ajuste de curvas de crescimento via modelos mistos na análise de dados longitudinais de genótipos de feijão vagem do programa de melhoramento genético de feijão vagem da Universidade Estadual Norte Fluminense Darcy Ribeiro – UENF.

2. REVISÃO DE LITERATURA

2.1. Feijão vagem

2.1.1 Origem

A origem evolutiva do gênero *Phaseolus* e a sua diversificação primária ocorreram nas Américas (Vavilov, 1931), mas o local exato onde isto se deu é ainda motivo de controvérsia (Gepts & Debouck, 1991). Não há um consenso sobre essas origens. Mas, há, no entanto, o senso comum entre os pesquisadores, de que realmente a origem e a domesticação ocorreram há mais de sete mil anos em dois centros de origem: a Meso-América – México e América Central – e a Região Andina (Kaplan, 1981).

A evolução e o melhoramento do feijão vagem ocorreram na França e nos Países Baixos, entre outros, sendo que as primeiras cultivares apropriadas para colheita e consumo das vagens ainda imaturas foram obtidas no início do século XIX, por meio de cruzamentos entre feijões cultivados na Europa e genótipos procedentes da América Central. Posteriormente, essas cultivares foram introduzidas na América do Norte, onde novos cruzamentos foram realizados com feijões da região, obtendo-se novas cultivares com maior potencial produtivo (Oca 1987). As primeiras cultivares de feijão vagem foram comercializadas nos Estados Unidos em 1887 e serviram como genitores para as cultivares modernas (Atkin, 1972).

2.1.2. Botânica e características

É uma hortaliça anual, herbácea, com sistema radicular superficial, do tipo pivotante e folhas trifolioladas. O caule pode ser uma haste nas cultivares de crescimento determinado ou volúvel nas cultivares de crescimento indeterminado. A coloração das flores pode variar conforme a cultivar, podendo ser encontradas flores brancas, rosa, roxa e lilás (Castellane et al., 1998; Portes, 1988; Krause et al., 2008; Filgueira, 2012; Silva et al., 2013).

Devido à estrutura de sua flor, o feijão vagem é uma planta autógama, que se multiplica por autofecundação, em geral com incidência muito baixa de polinização cruzada (Filgueira, 2008). A autogamia da espécie é favorecida pela estrutura de sua flor, na qual os órgãos masculinos e femininos estão bem protegidos na quilha, estrutura formada por 2 pétalas soldadas que envolvem o estigma e os estames em uma estrutura em forma de espiral, onde a polinização ocorre antes da abertura da flor. A fertilização ocorre de oito a nove horas depois da abertura floral (Vieira, 2006).

O fruto é do tipo legume, mais especificamente, uma vagem, que apresenta polpa espessa, formato afilado, podendo ser reta ou ligeiramente curvada, dentro das quais se desenvolvem as sementes (Filgueira, 2012). As sementes são um pouco mais compridas que as do feijão para grãos (Krause, 2008).

Segundo a Norma Brasileira 42 (NB42) (2012) as vagens podem ser de cor verde-escura, verde-clara, amarela, rajada e roxa. No entanto, as vagens verdes e amarelas são as mais comuns e de acordo com Castellane et al. (1988) e Abreu et al. (2004), ambas servem, tanto para consumo "*in natura*" quanto para a industrialização, sendo que as amarelas apresentam um mercado mais restrito. De forma geral o consumidor final prefere vagens de diâmetro menor, formato reto e com posição marginal do dente. Essas condições reduzem as perdas no processamento das vagens (Abreu, 2001; Souza, 2012).

Quanto ao teor de fibras, segundo Peixoto et al. (1993) se encontra entre 0,70 a 1,70%. Trabalhos desenvolvidos por Oliveira et al. (2001) e Oliveira et al. (2003) corroboram a faixa estabelecida. Porém, Abreu et al. (2004), encontraram teores de fibras nas vagens frescas no feijão vagem abaixo de 0,7%. Por outro lado Mariguele et al. (2008) obtiveram teor de fibra em torno de 4,78%.

Em relação ao hábito de crescimento, este pode ser determinado ou indeterminado. As cultivares de crescimento indeterminado são mais produtivas e exigem, para maior produção, o tutoramento das hastes, que podem atingir 2,5 m de comprimento (Abreu et al., 2000). O período de colheita inicia-se entre 50 e 80 dias a contar da semeadura e as colheitas são numerosas e frequentes (até três vezes por semana), prolongando-se por cerca de 30 dias, dependendo do estado nutricional e fitossanitário das plantas (Guedes, 2003; Filgueira, 2012).

As cultivares de crescimento determinado, também chamadas de arbustivas ou rasteiras ou anãs (Atházio, 1993), apresentam de 50 a 60 cm de altura (Portes, 1988; Barbosa et al., 2001) e a colheita inicia-se com 45 a 55 dias após semeadura, estendendo-se por duas a três semanas, podendo-se efetuar uma única colheita, com subsequente arranquio das plantas (Guedes, 2003). Nesse caso, o rendimento de vagens é relativamente menor, mas há a compensação do menor gasto com mão de obra (Pinto et al., 2001).

As vagens são colhidas imaturas, com sementes pouco desenvolvidas, sendo o ponto ideal de colheita quando as vagens atingirem o máximo desenvolvimento, porém antes de tornarem-se fibrosas. Na colheita, as vagens devem estar ainda tenras, com as extremidades podendo ser vergadas e quebradas sem esforço. A colheita no Brasil é manual e sua produtividade é variável (Barbosa et al., 2001; Guedes, 2003). Em condições ambientais favoráveis e com um manejo adequado, pode-se obter mais de 35 t.ha⁻¹ (Araújo et al., 2001).

Quanto à classificação, as plantas além de serem classificadas de acordo com o hábito de crescimento (determinado ou indeterminado) também podem ser divididas em tipos ou grupos. Segundo Filgueira (2012), podem ser encontrados os grupos macarrão e manteiga, de crescimento indeterminado e o grupo rasteiro de crescimento determinado, que possui vagens do tipo macarrão. Além destes, de acordo com a NB42 (2012) podem ser encontrados ainda os tipos holandesa e metro. Na CEASA-RJ as cultivares são classificadas somente em macarrão e manteiga, sendo a preferência pelo tipo manteiga.

As cultivares do grupo manteiga apresentam secção transversal achatada, perfil longitudinal semiarqueado, presença de fio nas fendas longitudinais entre as duas metades do fruto, ápice abrupto e dente apical arqueado marginal. As do tipo macarrão apresentam perfil elíptico, perfil longitudinal semiarqueado, ápice

abrupto, com dente apical semiarqueado marginal. As do tipo Holandesa apresentam formato elíptico, como a 'Macarrão', perfil longitudinal reto, ápice abrupto, com dente apical reto marginal. Já as do tipo Metro são bem mais compridas que as outras, com secção transversal circular, perfil longitudinal semiarqueado, ápice afilado e coloração verde intensa (NB42, 2012).

2.1.3. Importância econômica da cultura

O feijão vagem é um alimento consumido em diversos países, sendo estimado que a produção mundial de vagem esteja em torno de 6,5 milhões de t.ano⁻¹ (FAO, 2010), sendo que 250 a 300 mil toneladas deste montante sejam produzidas na América Latina por pequenos produtores que se utilizam de áreas de 2 a 20 ha, usando em sua maioria, germoplasma de porte indeterminado (Silbernagel et al., 1991; Henry & Jansen, 1992; Abreu et al., 2004).

No Brasil foram produzidas 56776 toneladas de vagem em 2006, sendo que 65,04% desse total foram produzidas pela região sudeste, com os estados de MG, SP e RJ, sendo responsáveis por 26,83%, 19,93% e 14,16% do total de vagens produzidas pelo país, respectivamente (IBGE, 2006).

No ano de 2012 foram comercializadas 5921,25 ton de vagem na CEASA-RJ, sendo 99,36% (5883,92 ton) do tipo manteiga e 0,64% (37,653 ton) do tipo macarrão. Do total comercializado, 98,65% (5841,59 ton) são provenientes do Estado do Rio de Janeiro, ou seja, apenas 1,35% (79,66 ton) foram importados de estados vizinhos. Toda a comercialização feita em 2012 com vagens de outros estados foi com o tipo manteiga, ou seja, não houve em 2012, importação de vagem do tipo macarrão (CEASA, 2012).

A região que mais contribuiu para a produção de feijão vagem do Estado do Rio de Janeiro no ano de 2012 foi a região serrana com 85,55% (4997,39 ton) da produção estadual, seguida da região centro sul fluminense com 11,59% da produção, ou seja, praticamente toda produção de feijão vagem do Estado do Rio de Janeiro é oriunda das regiões serrana e centro sul fluminense. A região Noroeste contribuiu com 1,51% da produção de feijão vagem.

A região Norte Fluminense contribuiu com 0,33% (18,975 ton) do total comercializado na CEASA-RJ em 2012, com os municípios, São João da Barra e

São Fidélis, sendo responsáveis pela comercialização de toda a produção da região. Do total comercializado pela região, São João da Barra – RJ, comercializou 77,47% (14,700 ton) e São Fidélis 22,53% (4,275 ton), sendo que toda produção da região foi representada pela vagem do tipo manteiga (CEASA, 2012).

Atualmente, a região Norte Fluminense contribui pouco para a produção de feijão vagem do estado. Entretanto, apresenta grande potencial de aumento de produção e produtividade. Nesse sentido, pesquisas têm sido realizadas na região com intuito de desenvolver novos cultivares sendo estes produtivos e adaptados às condições edafoclimáticas e socioeconômicas da região. A Universidade Estadual Norte Fluminense Darcy Ribeiro – UENF apresenta um papel importante nesse trabalho com o seu programa de melhoramento genético de feijão vagem de hábito de crescimento indeterminado.

2.1.4. Aspectos sobre a produção de feijão vagem

2.1.4.1. Clima

A vagem é uma cultura de larga adaptação a climas quentes e amenos, dentro de uma faixa térmica ampla (18°C a 30° C), sendo prejudicada por temperaturas acima de 35°C (Filgueira, 2012). Bohmer (2008) observou que em 11,88% do tempo, a temperatura do ar no interior do ambiente protegido esteve acima de 30 °C não causando nenhum prejuízo à emergência, ficando o tempo de emergência dentro do esperado. O mesmo comportamento foi observado por Aldrighi (2000). Na fase de intenso crescimento vegetativo o calor excessivo aumenta a fotorrespiração, reduzindo a taxa de crescimento (Mariot, 2000).

A fase reprodutiva é mais sensível a altas temperaturas (acima de 27°C), em que há deficiência de polinização, o que resulta em vagens deformadas (Barbosa et al., 2001; Omae et al., 2007). Hedlwein et al. (2010) observaram que a temperatura máxima diária do ar ultrapassou 32°C, em oito dias no ambiente protegido e em três dias no ambiente externo, antes dos estágios de início da floração e início de formação de vagens, sem serem limitantes para a cultura nos dois ambientes.

Por outro lado, a vagem é uma das hortaliças mais intolerantes ao frio e à geada (Filgueira, 2012). Baixas temperaturas reduzem o crescimento da planta, podendo induzir a formação de anomalias na frutificação, reduzir ou atrasar a germinação das sementes (Maroto, 2000; Constat, 2003) e aumentar o tempo de duração de cada estágio fenológico (Bohmer, 2008). Além disso, solos frios também promovem o apodrecimento das sementes e o frio também promove maior incidência de ferrugem (Peixoto et al., 2002).

Temperaturas abaixo de 12°C observadas por Aldrighi (2000) levaram a uma menor duração do tempo de colheita e a uma menor produtividade quando comparado a um cultivo de feijão vagem no período de primavera-verão.

O feijoeiro é uma planta sensível à deficiência hídrica e ao excesso de água no solo (Silveira & Stone, 2004). Seu requerimento de água varia com o estágio de desenvolvimento, de um valor mínimo na germinação até o máximo na floração e formação das vagens (Nóbrega et al., 2001). Nessa fase, a redução na produtividade sob estresse hídrico deve-se à baixa porcentagem de vingamento das flores, causando abortamento de óvulos, produzindo assim vagens chochas e queda de flores (Guimarães, 1998; Aidar et al., 2002).

Na fase vegetativa da cultura, a menor disponibilidade de água pode reduzir a taxa de crescimento da cultura (TCC), porque sendo a TCC o somatório das taxas de crescimento dos diversos componentes da planta, ou seja, da variação da matéria seca com o tempo, essa variável representa a capacidade de produção de fitomassa da cultura, isto é, sua produtividade primária (Benincasa, 1988).

Na falta de chuvas a irrigação é indispensável, sendo que, o volume de água a ser aplicado e o turno de rega vão depender das condições climáticas, do tipo de solo e do estágio de desenvolvimento da cultura (Almeida, 2011). Segundo Constat (2003), a cultura exige 200 a 300 mm de água por ciclo, sendo que o período entre a semeadura e a floração requer de 110 a 180 mm de água.

Trabalhando com feijão vagem do tipo Metro, Silva et al. (2012) não observam efeito do potencial matricial sobre o crescimento em altura da cultura. Por outro lado, observaram que as plantas que cresceram em solo mais úmido na fase vegetativa apresentaram maior diâmetro de colo, massa seca de folhas, massa seca de ramos e massa seca da parte aérea que aquelas que se desenvolveram em solo mais seco. Nesta condição a água fica retida mais

fortemente à superfície dos coloides; assim, as plantas precisam gastar mais energia para adquirir água para uso em seus processos fisiológicos vitais, fato que afeta negativamente as características de crescimento.

2.1.4.2. Época de plantio e cultivo protegido

Em localidades de inverno ameno é possível o plantio do feijão vagem ao longo do ano, possibilitando a comercialização na entressafra e cotações mais elevadas, especialmente em julho/agosto. Na maioria das regiões produtoras, a época normal de plantio é agosto/abril, evitando-se os meses de maio a julho, quando as temperaturas podem ser excessivamente baixas (Guedes, 2003). Porém, Amaro et al. (2007) recomendam para as Regiões Sudeste, Centro-Oeste, norte da Região Sul e sul do Nordeste, o plantio de agosto a fevereiro.

Comparando épocas de semeadura, Peixoto et al. (1997) observaram que o período de abril a junho foi o mais favorável para o cultivo tradicional de feijão vagem arbustivo em Goiás. Também em Goiás, Vidal et al. (2007) avaliando genótipos de feijão vagem arbustivo obtiveram maiores produtividades no cultivo de outono-inverno do que no primavera-verão devido à menor pressão de pragas e baixo estresse decorrente do calor.

Estudando a flutuação populacional de insetos associados ao feijão vagem, Menezes et al. (2003) observaram maiores populações de cigarrinha-verde nos plantios de março e maio. As maiores populações de mosca branca ocorreram no plantio de março, sendo que no plantio de maio praticamente não foram encontrados adultos. As populações de vaquinhas mantiveram-se baixas em quase todos os plantios estudados. A presença da cigarrinha-verde manteve-se constante, e em níveis relativamente altos, durante a maior parte das avaliações, demonstrando o papel de praga-chave da cultura. As maiores porcentagens de vagens com defeitos de encurvamento apareceram nos plantios de março e maio, coincidindo com altas populações de cigarrinha-verde (maio) e mosca-branca (março).

Em regiões em que as condições climáticas não possibilitam a implantação da cultura em determinadas épocas do ano pode-se cultivar o feijão vagem em ambiente protegido em que parte das limitações ambientais pode ser minimizada

(Heldwein et al., 2010). Além disso, possibilita o plantio em períodos de menor demanda, podendo aumentar os lucros do produtor (Basso et al., 2012).

O cultivo em ambiente protegido tende a proporcionar um ciclo mais longo comparado ao cultivo a campo nas mesmas condições, fato relacionado à menor disponibilidade de radiação solar dentro da estufa (Lopes et al., 1987; Bohmer, 2008; Heldwein et al., 2010). Souza & Escobedo (1997) observaram um aumento de 2 dias no ciclo do feijão vagem em cultivo protegido quando comparado com o cultivado em ambiente externo.

Apesar da menor disponibilidade de radiação solar no interior da estufa, o rendimento do feijão vagem foi similar nos dois ambientes estudados (externo e protegido), sugerindo que as plantas utilizam melhor a radiação no ambiente protegido pela estufa do que no ambiente externo, possivelmente por causa da maior disponibilidade de radiação solar difusa, a qual é mais eficiente em penetrar no interior do dossel vegetal (Heldwein et al., 2010).

Em uma condição de maior difusividade da radiação, a planta apresenta um maior índice de área foliar (Souza & Escobedo, 1997). Do estágio inicial de crescimento das culturas até a máxima interceptação da radiação pelas folhas, a TCC aumenta até um valor máximo, por causa do IAF (Bernardes, 1987), que por sua vez aumenta até as fases fenológicas R7- R8 (aparecimento da primeira flor e vagem, respectivamente) decrescendo nas fases seguintes, R8 e R9 (desenvolvimento das sementes e maturação, respectivamente) em que há senescência e queda de folhas no final do ciclo (Souza & Escobedo, 1997; Bohmer, 2008).

Em cultivo protegido ou a céu aberto o feijão vagem pode ser cultivado após o tomate ou pepino, por ser também uma cultura tutorada e não possuir custos adicionais de cultivo nesse sentido. O cultivo em sucessão é uma forma de utilizar os resíduos nutricionais presentes no solo, otimizar práticas agrícolas e o uso de mão de obra, bem como de poupar insumos, reduzindo os custos de produção (Peixoto et al., 2001; Lopes et al., 2004). Outra possibilidade relatada é o cultivo do feijão vagem de crescimento determinado após a batata.

Silva et al. (2001) mediante a análise química do solo antes e após o cultivo da batata observaram acréscimo de cálcio e fósforo e redução de Mg e K com as doses do adubo aplicadas no cultivo da batata. No plantio em sucessão do feijão vagem houve incremento no comprimento médio, número médio de

sementes por vagem e na produção total de vagem com a aplicação de 4 t ha⁻¹ de 4-16-8 realizada no cultivo da batata, não sendo necessária a adubação de plantio do feijão vagem cultivado após a cultura da batata.

2.1.4.3. Solo e adubação

De um modo geral, o feijão vagem desenvolve-se melhor em solos com textura média, não compactados, profundos, leves, com boa drenagem, com boa disponibilidade de água, ricos em matéria orgânica e férteis. Não tolera a acidez, produzindo melhor na faixa de pH 5,6 a 6,8 (Almeida, 2011). Em solos mais ácidos, a calagem deve ser efetuada, elevando-se a saturação por bases para 70%, procurando-se atingir pH 6,5 (Filgueira, 2012).

Quanto à adubação orgânica, diversas fontes de adubos podem ser utilizadas no cultivo do feijão vagem. Carrijo et al. (1999) recomendam a aplicação de 10 t.ha⁻¹ de esterco de curral curtido, sendo que no caso de rotação com o tomate ou pepino não há necessidade.

Alves et al. (2000) obtiveram a maior produção de sementes (3.555,10 kg.ha⁻¹) com a dose estimada de 27,66 t.ha⁻¹ e Santos et al. (2001) verificaram rendimento máximo de vagens (30,3 t.ha⁻¹) com a dose estimada de 24 t.ha⁻¹ de esterco bovino. Segundo estes autores, o esterco bovino, por apresentar maior rendimento e maior receita prevista, deve ser recomendado como adubo orgânico para esta hortaliça.

Trabalhando com esterco suíno, Araújo et al. (2001) observaram que o número, a produção de vagens por planta e a produtividade aumentaram de forma linear. A produtividade obtida com a maior dose (40 t.ha⁻¹) de esterco suíno foi de 35,4 t.ha⁻¹, sendo que os melhores resultados foram obtidos quando a adubação orgânica foi associada com a mineral. O emprego de 40 t.ha⁻¹ de esterco suíno aliado à adubação mineral proporcionou incremento de 4,2 t.ha⁻¹ de vagens, em relação ao emprego do esterco suíno sem adubação mineral.

Com esterco de galinha, Santos et al. (2001) obtiveram produtividade máxima (26,3 t.ha⁻¹) com a dose estimada de 13 t.ha⁻¹. Já, Oliveira et al. (2006) obtiveram produtividade máxima (26 t.ha⁻¹) da cultivar Alessa com a dose estimada de 20,3 t.ha⁻¹.

Com relação ao esterco caprino, Alves et al. (2000) obtiveram maior produção de sementes ($3.259,56 \text{ kg}\cdot\text{ha}^{-1}$) com a dose estimada de $20,85 \text{ t}\cdot\text{ha}^{-1}$ e Santos et al. (2001) obtiveram maior rendimento ($23 \text{ t}\cdot\text{ha}^{-1}$) com a dose de $17 \text{ t}\cdot\text{ha}^{-1}$. Quanto ao húmus de minhoca, os autores acima não encontram influência significativa sobre a produção de sementes e rendimento.

As adubações NPK de plantio e de cobertura podem ser feitas da seguinte forma: plantio – aplicar 30% do nitrogênio, 50% do potássio e todo o fósforo misturando junto com a adubação orgânica mais o solo antes da semeadura. Posteriormente em cobertura – parcelar em duas aplicações o restante do nitrogênio (70%) e do potássio (50%) (Carrijo et al., 1999).

Oliveira et al. (2003) trabalhando com produção de sementes da cultivar Macarrão Trepador, obtiveram produção máxima ($3.219 \text{ kg}\cdot\text{ha}^{-1}$) com a dose de $100 \text{ kg}\cdot\text{ha}^{-1}$ de sulfato de amônio. Utilizando ureia, a produção máxima ($2.221 \text{ kg}\cdot\text{ha}^{-1}$) foi estimada com a dose de $55 \text{ kg}\cdot\text{ha}^{-1}$ de N e com nitrato de cálcio a produção máxima ($2.517 \text{ kg}\cdot\text{ha}^{-1}$) foi estimada com a maior dose ($100 \text{ kg}\cdot\text{ha}^{-1}$). Provavelmente durante o crescimento e desenvolvimento das plantas, o nitrogênio fornecido, juntamente com os nutrientes contidos no solo, suprimiram eficientemente as necessidades nutricionais da cultura, permitindo inferir que o efeito das fontes e das doses de nitrogênio foi devido ao suprimento de nutrientes de forma equilibrada, conferindo à cultura a capacidade máxima de produção de sementes.

Por outro lado, Tavares Sobrinho (2001) obteve redução da produtividade do feijão vagem em função de doses elevadas de nitrogênio, fonte ureia. Doses acima de $55 \text{ kg}\cdot\text{ha}^{-1}$ de N, fonte ureia, também reduziram a produção de sementes do feijoeiro de acordo com Oliveira et al., (2003), possivelmente em consequência direta do efeito tóxico do amônio e da baixa taxa de nitrificação ou devido ao efeito indireto do amônio, reduzindo a absorção de outros cátions, isto é, exercendo forte efeito competitivo sobre os cátions (K^+ , Ca^{++} e Mg^{++}) de tal forma que a absorção destes seria reduzida (Carnicelli et al., 2000; Oliveira et al., 2003).

Nas regiões Sul e Sudeste, para o suprimento adequado de fósforo, pode-se aplicar de 200 a $300 \text{ kg}\cdot\text{ha}^{-1}$ de P_2O_5 em solos com baixo a médio teor deste elemento (Filgueira, 2012). Oliveira et al. (2005) trabalhando com a cultivar Macarrão Trepador obtiveram produção máxima de vagens ($30,13 \text{ t}\cdot\text{ha}^{-1}$) com a

dose de 252 kg.ha⁻¹ de P₂O₅, sendo que a dose capaz de proporcionar maior retorno econômico, foi 231 kg.ha⁻¹.

Em relação à adubação potássica, Oliveira et al. (2007) observaram influência positiva do K sobre o número de vagens, produção por planta e a produtividade. A produtividade de vagens atingiu valor máximo estimado de 25,3 t.ha⁻¹ na dose de 171 kg.ha⁻¹ de K₂O. Filgueira (2012) recomenda a aplicação de 60 a 80 kg.ha⁻¹ de K₂O.

2.2.1. Dados longitudinais

Na pesquisa com experimentos, a metodologia de análise estatística dos dados é selecionada de acordo com a estrutura destes, a qual poderá apresentar-se na estrutura univariada ou multivariada. A estrutura univariada é obtida quando para cada unidade experimental somente uma variável resposta é avaliada. Quando esta estrutura é multivariada tem-se a avaliação de mais de uma variável resposta em uma mesma unidade experimental. Outra forma de estrutura multivariada é obtida quando para a mesma unidade experimental avalia-se uma variável, em diferentes condições de avaliação, constituindo assim a estrutura de dados denominada de dados longitudinais (Amador, 2010).

Estudos envolvendo dados longitudinais desempenham um papel fundamental nas mais variadas áreas do conhecimento, recebendo denominações diferentes em outros campos da ciência. Segundo Singer et al. (2012), em Bioestatística, os dados longitudinais são conhecidos como *coorte*, ao passo que em outros campos do conhecimento, como Sociologia, Economia ou Administração, como *painel*. Esses autores discutem ainda que estudos longitudinais diferem de estudos conhecidos como *séries de tempo* ou *séries cronológicas* na medida em que nestes, em geral, uma única unidade amostral é avaliada em muitos instantes (200 ou mais, por exemplo), enquanto que em estudos longitudinais as unidades amostrais são observadas em poucas ocasiões.

Os estudos longitudinais constituem um caso especial dos estudos de medidas repetidas, que abrange o esquema conhecido como parcelas subdivididas e *crossover* (Costa, 2003). O termo medidas repetidas é usado para

designar medidas feitas na mesma unidade experimental em mais de uma ocasião (Diggle, 1988; Crowder & Hand, 1990; Littell et al., 1996). Os planejamentos com medidas repetidas do tipo “split - plot” (parcelas subdivididas) são comuns nas ciências agrárias, em que um único nível de um fator é alocado a uma parcela relativamente grande de terra e todos os níveis de um segundo fator, são aplicados às subparcelas. Nos planejamentos do tipo “crossover”, as unidades experimentais recebem seqüências de tratamentos (Freitas, 2007).

O estudo básico de medidas repetidas consiste de delineamentos experimentais completamente aleatorizados em que os tratamentos são alocados às unidades experimentais e os dados são coletados mais de uma vez em cada unidade experimental. Dessa forma, têm-se pelo menos dois fatores a serem estudados (tratamento e tempo), em que tratamento é o fator entre indivíduos e tempo, o fator intra-indivíduos (Costa, 2003; Rosário et al., 2005). Rezende et al (1999) complementam que a interação época x tratamento é, provavelmente, o mais importante aspecto destes planejamentos, e quando significativa, indica uma tendência não paralela da resposta em relação ao tempo para diferentes níveis de tratamentos.

De forma geral os estudos envolvendo dados longitudinais têm como objetivos o estudo do comportamento da variável resposta ao longo do tempo e a verificação da existência de dependência da variável resposta em relação às covariáveis (Wald, 2000; Costa, 2003; Amador, 2010). No entanto, Riboldi (1999) relata os seguintes objetivos:

a) comparação das diferentes subpopulações quanto ao padrão de variação das respectivas distribuições de respostas ao longo das diferentes condições de avaliação, isto é, verificar se os perfis são paralelos, ou seja, a interação entre os tratamentos e o tempo é nula;

b) comparação das diferentes subpopulações quanto às respectivas distribuições das médias (em relação às diversas condições de avaliação) de respostas, isto é verificar se os perfis são coincidentes, ou seja, verificação de efeitos de tratamentos;

c) comparação das diferentes subpopulações quanto às respectivas distribuições das médias (em relação às diversas subpopulações) de respostas, isto é, verificar se os perfis são horizontais, ou seja, verificação de efeitos das condições de avaliação;

d) ajuste de modelos para explicar a variação das respostas médias como função do fator que define as condições de avaliação (ajuste de curvas).

Como vantagens, Venezuela (2003) cita os seguintes fatos:

- requerem menos unidades experimentais do que planejamentos completamente casualizados;

- proporcionam condições mais adequadas para o estudo de covariáveis que possam ter influencia na variável resposta;

- melhoram em geral, a precisão dos contrastes associados às diferenças entre os valores médios da variável resposta;

- permitem o estudo da mudança de comportamento da resposta média da unidade experimental nos diferentes tratamentos (incorpora informação sobre a variação individual da análise).

Por outro lado a maior desvantagem dos estudos longitudinais está relacionada com seu custo, pois em muitas situações exige-se um grande esforço para garantir a observação das unidades amostrais nos instantes predeterminados e em outras, o período de observação pode ser muito longo. Os aspectos técnicos também podem ser considerados como uma desvantagem, pois a análise estatística de dados obtidos sob esse tipo de planejamento é, em geral, mais complexa que a análise de dados obtidos sob esquemas transversais. (Singer et al., 2012).

A característica distinta de estudos longitudinais é a dimensão ordenada com que os dados são coletados e o fato de que as observações repetidas para um indivíduo tendem a ser correlacionadas. Tal correlação pode ser modelada

através de uma estrutura de covariância dos dados observados sendo que, para outros tipos de dados, é usual assumir que os erros sejam independentes. Modelar uma estrutura de covariâncias apropriada é essencial para que as inferências sobre as médias sejam válidas (Costa, 2003).

Neste tipo de análise, a variável resposta pode ser contínua ou discreta, avaliada nas diferentes unidades experimentais (indivíduo, vasos, canteiros), que podem ser agrupadas segundo tratamentos ou fatores. Para cada unidade experimental obtêm-se diversas unidades de observação, que em conjunto, definem um perfil individual de respostas. Para cada tratamento (ou grupo) está associado um perfil médio de respostas que deve evidenciar o efeito do tratamento e o seu comportamento ao longo do tempo (Freitas, 2007).

Os dados associados a cada unidade experimental podem ser expressos em termos de um vetor, com valores observados das variáveis respostas em cada instante e, de uma matriz, com os correspondentes valores das variáveis explanatórias ou covariáveis. O vetor com as p_i respostas da i -ésima unidade experimental ($i = 1, \dots, n$), chamado de perfil individual, pode ser expresso como $Y_i = (y_{i1}, \dots, y_{ip_i})^t$, em que y_{ik} representa o valor observado da variável resposta da i -ésima unidade experimental no t -ésimo instante de observação ($t = 1, \dots, p_i$). A matriz de variáveis explanatórias, de dimensões $p_i \times m$ pode ser expressa como $X_i = (x_{i1}, \dots, x_{im})$, em que: $x_{ij} = (x_{ij1}, \dots, x_{ijp_m})^t$; com x_{ijt} , o valor da j -ésima variável explanatória ($j = 1, \dots, m$) para a i -ésima unidade experimental no t -ésimo instante de observação (Araújo, 2007).

Devido à natureza dos estudos longitudinais, que podem durar meses ou até anos, e requerem observações nas mesmas unidades amostrais em múltiplas ocasiões, torna-se frequente a existência de dados incompletos, não balanceados, e irregulares no tempo (Littell et al., 2006).

Os dados longitudinais são denominados regulares no tempo se o intervalo entre duas medidas consecutivas quaisquer for constante ao longo do estudo, caso contrário, serão irregulares no tempo (Suyama, 1995). Se as observações forem feitas nos mesmos instantes de tempo em todas as unidades experimentais, ter-se-á uma estrutura de dados balanceada. Quando houver observações perdidas, então a estrutura dos dados é incompleta e se não houver observações perdidas, então a estrutura dos dados é completa (Freitas, 2007; Amador, 2010).

De acordo com Lima (1996), a análise de dados longitudinais pode ser realizada através da análise univariada de perfis, análise multivariada de perfis e através da análise de curvas de crescimento (que também pode ser realizada através de modelos uni e multivariado e modelos mistos).

Segundo Aubin (1984), a análise de perfil é uma técnica clássica de análise de dados longitudinais que pode ser realizada utilizando-se técnicas multivariadas ou univariadas e tem como principal objetivo testar hipóteses sobre os valores médios da variável resposta nas diversas ocasiões de observação e compará-los entre os diferentes grupos de unidades experimentais, ou tratamentos.

Quando o objetivo da pesquisa é estudar o crescimento, o modelo matemático usado nessas situações consiste na descrição do comportamento da variável em função do tempo, através de uma curva, geralmente de forma polinomial, em que as comparações entre efeitos de tratamentos diferentes podem ser estudadas através de comparações dos parâmetros das respectivas curvas (Singer, 1977; Castro, 1999), isto é, a diferença entre dois tratamentos quaisquer pode ser avaliada entre os dois coeficientes angulares e lineares. Além disso, o ajuste de curvas possibilita fazer previsões para o comportamento médio das respostas ao longo do tempo, utilizando, em geral, poucos parâmetros (Amador, 2010).

Lima (1988) destaca algumas vantagens no ajuste de curvas:

- possibilitam que sejam feitas previsões sobre o comportamento médio da resposta, para ocasiões diferentes daquelas para as quais o estudo foi planejado;
- possibilitam que sejam feitas comparações mais sensíveis entre as subpopulações, em função do número reduzido de parâmetros;
- possibilitam a aquisição de um maior conhecimento sobre o processo de variação temporal da resposta.

A seguir será apresentada uma revisão sobre algumas metodologias para análise de dados longitudinais. O enfoque da revisão será direcionado para a utilização da análise univariada de perfil e a técnica de modelos mistos.

2.2.2. Análise univariada de perfil

A análise univariada de perfil é uma das técnicas mais difundidas na análise de dados longitudinais, baseia-se no número total de observações e requer algumas pressuposições na estrutura de variâncias e covariância entre as observações. Esta técnica utiliza o esquema de parcelas subdivididas, em que os dados longitudinais representam as observações tomadas em subdivisões das parcelas, considerando os tratamentos e o fator tempo, como efeito de parcela e subparcela, respectivamente. Sendo assim, os tratamentos e o tempo (as condições de avaliação) são considerados como dois fatores fixos e cruzados e, as unidades experimentais, como um fator aleatório hierárquico dentro dos tratamentos (Amador, 2010).

Para o modelo univariado, no esquema de parcelas subdivididas, são feitas pressuposições de que, tanto o erro da parcela (erro entre indivíduos), que engloba o fator tratamentos ou grupos, como o erro da subparcela (erro intra-indivíduos), onde são alocados os tempos e a interação tratamentos x tempos, tenham distribuição normal, sejam independentes e identicamente distribuídos, com variâncias constantes (Xavier e Santos, 2001; Freitas, 2007).

Na análise de variância, considerando-se que o experimento foi instalado seguindo o esquema de parcelas subdivididas no tempo (análise de medidas repetidas), tem-se o seguinte modelo matemático sugerido por Vonesh & Chinchilli (1997):

$$Y_{ijk} = \mu + \beta_j + \tau_i + (\beta\tau)_{ji} + \gamma_k + (\tau\gamma)_{ik} + \varepsilon_{ijk},$$

em que:

Y_{ijk} : é o valor observado para a variável resposta no i -ésimo tratamento para o j -ésimo bloco no k -ésimo tempo;

μ : é uma constante inerente a todas as observações;

β_j : é o efeito do j -ésimo bloco;

τ_i : é o efeito do i-ésimo tratamento;

$(\beta\tau)_{ji}$: é o efeito aleatório da interação do i-ésimo tratamento com o j-ésimo bloco (erro associado às parcelas);

γ_k : é o efeito do k-ésimo tempo observado;

$(\tau\gamma)_{ik}$: é o efeito da interação entre o i-ésimo tratamento com o k-ésimo tempo;

\mathcal{E}_{ijk} : é o erro aleatório correspondente às observações do k-ésimo tempo para o i-ésimo tratamento no j-ésimo bloco (variação do acaso sobre as observações), supostos homocedásticos, independentes e normalmente distribuídos.

O esquema de análise de variância para os fatores de interesse do modelo univariado (x) será o representado na tabela 1.

Tabela 1. Análise de variância, Esperanças dos Quadrados Médios e Teste F.

Causas de Variação	G.L.	Q.M	E (Q.M.)	F
<i>Bloco (B)</i>	$(b-1)$	$Q M_1$	$\sigma^2 + t\sigma_{bg}^2 + gt\sigma_b^2$	
<i>Tratamento (G)</i>	$(g-1)$	$Q M_2$	$\sigma^2 + t\sigma_{bg}^2 + bt\phi_g$	$F_1 = Q M_2/Q M_3$
<i>Erro (a) (BG)</i>	$(b-1)(g-1)$	$Q M_3$	$\sigma^2 + t\sigma_{bg}^2$	
<i>Tempo (T)</i>	$(t-1)$	$Q M_4$	$\sigma^2 + gb\phi_t$	$F_2 = Q M_4/Q M_6$
<i>Trat x Tempo (GT)</i>	$(g-1)(t-1)$	$Q M_5$	$\sigma^2 + b\phi_{gt}$	$F_3 = Q M_5/Q M_6$
<i>Erro (b) (BGT)</i>	$g(b-1)(t-1)$	$Q M_6$	σ^2	
<i>Total Cor.</i>	$gbt-1$			

Onde:

$$\phi_g = \frac{\sum \tau_i^2}{g-1}, \quad \phi_t = \frac{\sum \gamma_k^2}{t-1}, \quad \phi_{gt} = \frac{\sum (\tau\beta)_{ik}^2}{(g-1)(t-1)}$$

Considerando o modelo univariado, no esquema de parcelas subdivididas será testado se existe diferença entre tratamentos, entre as condições de avaliação (tempo) e entre as interações entre tratamentos x tempos, por meio das estatísticas F_1 , F_2 e F_3 que seguem distribuições exatas F centrais dadas pelos graus de liberdade apresentados na tabela 1 (Steel; Torrie, 1988; Lima, 1996). Desta forma, as hipóteses de interesse a serem testadas são:

a) a estatística F_1 , testa a hipótese:

$$H_0 : \tau_1 = \tau_2 = \tau_3 = \dots = \tau_g = 0$$

$$H_1 = \text{pelo menos um } \tau_i \neq 0$$

Quando H_0 é rejeitada, **indica que os perfis não são coincidentes**, ou seja, existem diferenças verticais entre os tratamentos;

b) a estatística F_2 , testa a hipótese:

$$H_0 = \gamma_1 = \gamma_2 = \gamma_3 = \dots \gamma_t = 0$$

$$H_1 = \text{pelo menos um } \gamma_k \neq 0$$

Quando H_0 é rejeitada, **indica que os perfis não são horizontais**, ou seja, as curvas não são constantes ao longo do tempo;

c) a estatística F_3 , testa a hipótese:

$$H_0 = (\tau\gamma)_{11} = (\tau\gamma)_{12} = (\tau\gamma)_{ik} = 0$$

$$H_1 = \text{pelo menos um } (\tau\beta)_{ik} \neq 0$$

Quando H_0 é rejeitada, **indica que os perfis não são paralelos, ou seja**, a forma das curvas difere entre os tratamentos ao longo do tempo.

Conforme visto anteriormente, a estrutura de parcelas subdivididas, em um estudo de medidas repetidas no tempo, é caracterizada quando se aplicam às parcelas os níveis do fator A e nos quais se tomam medidas repetidas, em ocasiões sucessivas, sob a mesma parcela, admitindo-se que essas medidas, tomadas em ocasiões distintas, têm variâncias homogêneas e são igualmente correlacionadas. No entanto, os níveis desse tempo não podem ser aleatorizados para os seus intervalos (Xavier e Santos, 2001).

A falta de aleatorização para os níveis do fator tempo, em um esquema de parcelas subdivididas no tempo, implica que os erros correspondentes às respectivas unidades experimentais ou indivíduos podem ter uma matriz de covariâncias que não é aquela exigida para que a análise usual de um delineamento seja válida, isto é, variâncias homogêneas (Xavier e Santos, 2001). Assim, o uso do esquema em parcelas subdivididas no tempo, que corresponde a uma análise univariada dos perfis, nem sempre é recomendado para análise de dados com medidas repetidas, porque pressupõe que a estrutura da matriz de covariância seja do tipo uniforme, (variâncias iguais nas diversas ocasiões e covariâncias iguais entre duas ocasiões quaisquer) (Barbosa, 2009). Além disso, quando o experimento é sistematicamente arranjado, sem aleatorização, a análise de um experimento de medidas repetidas com esquema de parcelas subdivididas pode aumentar a probabilidade de falsamente rejeitar a hipótese nula (erro tipo I) (Fernandez, 1991).

Em experimentos com medidas repetidas, quando analisados sob o enfoque de parcela subdividida, as medidas repetidas em uma mesma unidade experimental ao longo do tempo são, no geral, correlacionadas (Malheiros, 1999). Como consequência imediata de se ignorar a correlação entre os dados mensurados em tempos adjacentes tem-se que a significância aparente da diferença entre as medias dos tratamentos é grosseiramente exagerada e a sensibilidade dos testes para interação é seriamente reduzida. Quando a correlação de erros é ignorada, as inferências podem ser distorcidas, conforme o grau de homogeneidade das variâncias e covariâncias dos dados nas diferentes épocas (Gill, 1986; Riboldi et al., 1996).

A análise univariada é válida para testes de efeitos de tratamentos independentes da uniformidade da matriz de variâncias e variância-covariância, ou seja, se aceita a pressuposição de homocedasticidade, e a estatística F_2 pode

ser utilizada mesmo que a condição de uniformidade não seja satisfeita. A condição de uniformidade interfere apenas nas estatísticas F_2 e F_3 (Milliken e Johnson, 1992; Castro, 1997). Uma condição suficiente, mas não necessária para que as estatísticas F_2 e F_3 , ou seja, teste F em nível de subparcela para o fator tempo e interação tempo x tratamento, tenham distribuição exata F central é que a matriz de variância e covariância tenha a forma de simetria composta (Castro, 1999; Xavier e Santos, 2001).

A condição de simetria composta implica que a variável aleatória seja igualmente correlacionada e tenha variâncias iguais, considerando as diferentes ocasiões. Porém, as matrizes de covariâncias, na forma da simetria composta e erros independentes, são casos especiais de uma condição mais geral para a matriz de variância e variância-covariância, denominada de HUYNH-FELDT (H-F) ou condição de esfericidade (circularidade) (Xavier e Santos, 2001). A condição H-F é equivalente a especificar que as variâncias das diferenças entre pares de erros sejam todas iguais. As variâncias das medidas feitas nas mesmas ocasiões sendo iguais são equivalentes à simetria composta (Freitas, 2007).

A condição de H-F é uma condição necessária e suficiente para que o teste F da análise de variância usual, no esquema de parcelas subdivididas no tempo, seja válido. A condição H-F específica para que os elementos da matriz de covariâncias sejam expressos, para um $\lambda > 0$, como:

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \frac{(\sigma_1^2 + \sigma_2^2)}{2} - \lambda & \frac{(\sigma_1^2 + \sigma_3^2)}{2} - \lambda & \frac{(\sigma_1^2 + \sigma_4^2)}{2} - \lambda \\ \frac{(\sigma_1^2 + \sigma_2^2)}{2} - \lambda & \sigma_2^2 & \frac{(\sigma_2^2 + \sigma_3^2)}{2} - \lambda & \frac{(\sigma_2^2 + \sigma_4^2)}{2} - \lambda \\ \frac{(\sigma_3^2 + \sigma_1^2)}{2} - \lambda & \frac{(\sigma_3^2 + \sigma_2^2)}{2} - \lambda & \sigma_3^2 & \frac{(\sigma_3^2 + \sigma_4^2)}{2} - \lambda \\ \frac{(\sigma_4^2 + \sigma_1^2)}{2} - \lambda & \frac{(\sigma_4^2 + \sigma_2^2)}{2} - \lambda & \frac{(\sigma_4^2 + \sigma_3^2)}{2} - \lambda & \sigma_4^2 \end{bmatrix}$$

em que λ é a diferença entre a média das variâncias e a média das covariâncias.

Para testar se a matriz de covariâncias atende a condição H-F, ou seja, a condição de circularidade ou esfericidade pode-se utilizar o teste de esfericidade de Mauchly (1940), que verifica se uma população normal multivariada apresenta variâncias iguais e correlações nulas. Caso satisfaça essa condição, a matriz de covariâncias será chamada de esférica e implicará na garantia que os testes F da

subparcela, no caso da parcela subdividida no tempo possuirá distribuição F exata.

O teste de Mauchy é bastante sensível a não normalidade dos dados (Crowder e Hand, 1990). Não é muito poderoso para amostras pequenas e não é robusto quando há violação da suposição de normalidade (Vonesh e Chinchilli, 1997).

Se a condição de esfericidade da matriz de covariâncias for satisfeita (ou seja, se o teste de Mauchly resultar não significativo), pode-se analisar os dados utilizando-se a técnica de análise univariada de perfis, admitindo um modelo de parcelas subdivididas no tempo. Se a condição de esfericidade da matriz de covariâncias não for satisfeita (ou seja, se o teste de Mauchly resultar significativo) uma alternativa univariada consiste em corrigir os graus de liberdade das estatísticas dos testes envolvendo comparações intra-indivíduos, realizando uma análise univariada aproximada (Barbosa, 2009).

Fernandes (1991) sugere que a correção para os números de graus de liberdade deve ser realizada somente quando a condição H-F (circularidade ou esfericidade) para a matriz de covariâncias não for satisfeita, e o nível de significância do teste de esfericidade estiver entre 0,05 e 0,01.

2.2.2.1 Correção dos graus de liberdade

As correções para os números de graus de liberdade foram inicialmente propostas por Box (1954a, 1954b) e, posteriormente, aperfeiçoadas por Geisser & Greenhouse (1958) e Huynh & Feldt (1976). Todas são fundamentadas na multiplicação do número de graus de liberdade das fontes de variação associados às subparcelas por valores estimados para cada método. Por exemplo, dada uma estatística F, baseada em mínimos quadrados, com números de graus de liberdade ν_1 e ν_2 , o ajuste para esses números de graus de liberdade seria: $\varepsilon \cdot \nu_1$ e $\varepsilon \cdot \nu_2$; com ε estimado por um dos métodos (essa notação é variável entre autores) (Silva, 2011). As correções devem ser feitas apenas em estatísticas que envolvam comparações “dentro”, ou seja, intra-indivíduos. Portanto, a estatística F_1 não necessita de correção de seu número de grau de liberdade, qualquer que seja a estrutura de covariância. Está associada apenas a comparações “entre”, ou

seja, inter-indivíduos; F_2 sempre terá distribuição exata F central. As estatísticas F_3 e F_4 podem precisar de correções, pois envolvem comparações intra-indivíduos, que estão associadas diretamente a estrutura da matriz (Freitas, 2007). Como exemplos de ajustes para o número de graus de liberdade do teste F para o fator erro intra-indivíduos pode-se citar os ajustes de Geisser e Greenhouse (1958) e o ajuste de Huynh e Feldt (1976).

Box (1954a) estabeleceu uma medida do afastamento em relação à condição H-F, denotada por θ , que varia no intervalo $[(t-1)^{-1}; 1,0]$; assim, quanto mais θ se afasta de 1,0, mais a estrutura de Σ se desvia da condição de esfericidade (H-F), e maior será a redução sobre ν_1 e ν_2 . Notadamente, valores iguais ou próximos à unidade indicam que a análise pode ser realizada como se o experimento tivesse sido planejado em parcelas divididas; e se t (número de medidas repetidas) for igual a dois, a condição H-F é sempre satisfeita e o experimento pode ser analisado como parcelas divididas (ex. experimentos varietais em cana-de-açúcar com apenas dois cortes). Na prática, a correção de Box resume-se em dividir os graus de liberdade das fontes de variação tempo, interação, tratamento x tempo e do erro(b), por $(t-1)$, reavaliando-se as significâncias do teste F (Milliken & Johnson, 1992; Vivaldi, 1999).

As correções propostas por Geisser & Greenhouse (G-G) e Huynh & Feldt (H-F) também são efetuadas mediante ajuste nos graus de liberdade, com reflexos nos níveis de significância observados para a estatística F; dando-se, igualmente, pela multiplicação de valores estimados (denotados por $\hat{\epsilon}$ e $\tilde{\epsilon}$, epsilons; ou, respectivamente, θ_{G-G} e θ_{H-F}), aos números de graus de liberdade de fatores associados às subparcelas. Do mesmo modo, valores iguais ou próximos de 1,0 indicam que a análise pode ser realizada como se o experimento tivesse sido planejado em parcelas divididas. Os estimadores de $\hat{\epsilon}$ e $\tilde{\epsilon}$ são descritos em Xavier (2000) e Freitas (2007).

Outra alternativa para dados em que a condição de esfericidade da matriz de covariâncias não for satisfeita (ou seja, se o teste de Mauchly resultar significativo) é a realização de uma análise multivariada de perfis.

2.2.3. Análise multivariada de perfil

A denominação de análise multivariada de perfil deve-se ao fato de serem analisadas simultaneamente as relações entre as medidas múltiplas temporais da variável tempo para cada indivíduo (unidade experimental), com alguma outra variável, usualmente denominada de tratamento (Ferreira, 2012).

Segundo Vonesh & Chinchilli (1997), geralmente, as técnicas usuais impõem a suposição de que todas as observações sejam independentes, mas essa suposição não é adequada para dados de medidas repetidas onde as observações feitas no mesmo indivíduo usualmente são correlacionadas. Para um experimento com esquema de parcelas subdivididas, com medidas repetidas no tempo utiliza-se o seguinte modelo:

$$Y_{ijk} = \mu + \beta_j + \tau_i + (\beta\tau)_{ji} + \gamma_k + (\tau\gamma)_{ik} + \varepsilon_{ijk},$$

em que:

Y_{ijk} : é o valor observado para a variável resposta no i-ésimo tratamento para o j-ésimo bloco no k-ésimo tempo;

μ : é uma constante inerente a todas as observações;

β_j : é o efeito do j-ésimo bloco;

τ_i : é o efeito do i-ésimo tratamento;

$(\beta\tau)_{ji}$: é o efeito aleatório da interação do i-ésimo tratamento com o j-ésimo bloco (erro associado às parcelas);

γ_k : é o efeito do k-ésimo tempo observado;

$(\tau\gamma)_{ik}$: é o efeito da interação entre o i-ésimo tratamento com o k-ésimo tempo;

\mathcal{E}_{ijk} : é o erro aleatório correspondente às observações do k-ésimo tempo para o i-ésimo tratamento no j-ésimo bloco (variação do acaso sobre as observações), supostos homocedásticos, independentes e normalmente distribuídos.

Várias estatísticas de teste disponíveis são obtidas através dos princípios da união-intersecção de Roy e da razão de verossimilhança de Wilks. Definindo-se $\theta_i = \lambda_i (1 + \lambda_i)^{-1}$, em que λ_i é a i-ésima raiz característica (autovalor) de \mathbf{HE}^{-1} , sendo: \mathbf{H} matriz da soma de quadrados e produtos cruzados devido à hipótese nula; e \mathbf{E} a matriz de soma de quadrados e produtos cruzados devido ao erro (Xavier, 2000). As estatísticas são dadas por:

- i) Pillai's Trace: $P = \sum \theta_i$, com $i = 1, 2, \dots, s$; sendo $s = \min(t-1, g-1)$; em que, 't' é o número de tempos; e 'g', o de tratamentos;
- ii) Wilks' Lambda: $\Delta = \prod (1 - \theta_i)$;
- iii) Hotelling – Lawley Trace: $T = \sum \theta_i (1 - \theta_i)^{-1}$;
- iv) Roy's Greatest Root: $\theta_s = \max(\theta_i)$.

As distribuições exatas dessas estatísticas, sob a hipótese nula, dependem unicamente dos parâmetros $m_1 = (|(t-1) - (g-1)| - 1) / 2$ e $m_2 = (n-g - (t-1)-1) / 2$. Os testes estatísticos associados à matriz de covariância (Σ) do modelo multivariado não consideram a condição H-F. Além disso, para obter o número de parâmetros da matriz (Σ) basta utilizar a expressão: $t(t+1) / 2$, em que "t" corresponde ao número de medidas repetidas no tempo. As hipóteses testadas na análise multivariada foram as mesmas da análise univariada de perfil.

A análise multivariada não impõe nenhuma restrição ou condição para a matriz de covariância dos efeitos. A matriz de variâncias e covariâncias é chamada não estruturada ou completamente parametrizada. Meredith e Stehrnan (1991) também ressaltam que não há suposição sobre a estrutura da matriz de covariâncias, sendo, por isso, uma solução natural para dados de medidas repetidas. Por outro lado, Lima (1996) destaca que a análise multivariada requer a necessidade de perfis individuais de resposta completos e apresenta baixo poder dos testes de hipóteses, devido à excessiva parametrização.

Outra restrição para a utilização da análise multivariada é que a quantidade de parâmetros associados à matriz não estruturada, que precisa ser estimada, é

proporcional ao número de condições de avaliação. Em situações em que o número é grande (quando o número de unidades experimentais é pequeno em relação ao número de ocasiões de avaliação ou quando há presença de muitas observações incompletas), a eficiência dos estimadores dos parâmetros pode ser afetada. Segundo Wald (2000) em alguns casos é impossível estimar os parâmetros dessa matriz de covariância.

Perri et al. (1999), afirmam que uma abordagem mais atual consiste no uso de modelos lineares mistos, que se baseia em três aspectos fundamentais: estimação e teste de hipóteses sobre os parâmetros de efeito fixo, predição dos parâmetros de efeito aleatório e estimação dos componentes de variância e que segundo Camarinha Filho (2002), o sucesso do procedimento de modelagem está fortemente associado ao exame dos efeitos aleatórios e na possibilidade de se introduzir, no modelo, estruturas de variâncias e covariâncias.

2.2.4. Modelos Mistos

Em um modelo matemático, deseja-se explicar as variações de uma variável dependente por meio dos efeitos diferenciais que se atribuem a outra série de variáveis independentes. Tais efeitos podem ser de natureza fixa ou aleatória, conforme representem, respectivamente, constantes a serem estimadas ou realizações de uma variável aleatória com distribuição de probabilidade conhecida (Jomar, 2002). Considerando a natureza dos efeitos como fixos ou aleatórios os modelos estatísticos podem ser classificados em fixos, aleatórios ou mistos.

Segundo Searle (1987), modelos lineares nos parâmetros possuem ao menos um efeito aleatório (comumente denotado por erro experimental). Se um modelo apresenta todos os demais componentes fixos é chamado de modelo fixo; se, no entanto, todos os demais fatores forem aleatórios o modelo é chamado de aleatório; quando o modelo apresenta tanto efeitos aleatórios como fixos, é denominado de modelo misto. Estes últimos são usados para modelar a parte aleatória pela inclusão de uma matriz de variância-covariância (Littell et al., 2006).

Os modelos lineares mistos têm sido um tema de crescente interesse em Estatística nos últimos cinquenta anos, pois possibilitam a modelagem de

correlação intra-indivíduo, muitas vezes presente em dados agrupados. Observações feitas no mesmo indivíduo não podem ser consideradas não correlacionadas e os modelos lineares mistos constituem uma ferramenta conveniente para modelar a dependência intra-indivíduos (Pinheiro, 1994), motivo pelo qual torna os modelos mistos mais flexíveis do que os modelos univariados (análise sob o esquema de parcela subdividida no tempo) e multivariados.

A metodologia de modelos mistos foi inicialmente, proposta por Henderson (1949, 1975) e tornou-se conhecida nas últimas décadas devido ao desenvolvimento de suportes computacionais robustos (Amador, 2010). Antes desses avanços, as análises de modelos mistos eram executadas adaptando-se métodos para modelos de efeitos fixos. Isso trazia limitações às análises realizadas pelo proc GLM no SAS. As versões recentes do SAS incluem o proc mixed, que permitem a modelagem da estrutura de variância-covariância dos dados e o cálculo de estimativas eficientes dos efeitos fixos e de seus respectivos erros padrão (Littell, 2006).

O modelo misto simplifica e unifica diversas análises estatísticas comuns. Além da análise de dados espaciais e de modelos de efeitos aleatórios, o modelo misto é útil para: análise de planejamento de experimentos longitudinais; planejamento em parcelas subdivididas; meta análise e modelos de variância heterogênea. Além disso, extensões úteis de modelo misto estão disponíveis para modelos lineares generalizados e modelos não lineares (Pascual, 2000). O modelo misto também é útil para ajuste de curvas de crescimento (Lima, 1996), pois permite a inclusão de efeitos fixos e aleatórios no processo de modelagem e apresenta as seguintes vantagens (Laird & Ware, 1982):

- permitem a modelagem da matriz de variância-covariância com um número menor de parâmetros que o de modelos multivariados gerais;
- podem ser usados quando os dados não são balanceados em relação ao tempo;
- possibilitam o ajuste de curvas polinomiais de graus diferentes para cada grupo de unidades experimentais.

Nos modelos de curvas de crescimento, o interesse básico está na estimação e na predição dos efeitos de tratamentos em algum tempo, por

exemplo. Diferenciando-se do modelo usual de medidas repetidas no tempo, no qual o interesse é diagnosticar diferenças nos efeitos dos tratamentos ao longo do tempo (Ferreira, 2012).

Para análise de dados longitudinais existem dois modelos mistos úteis: o modelo padrão de covariância e o modelo de coeficientes aleatórios. O modelo padrão de covariância é utilizado quando os dados são balanceados; em contrapartida, o modelo de coeficientes aleatórios é apropriado para dados desbalanceados ou se há interesse particular na relação entre a variável resposta e o tempo, caso típico das curvas de crescimento. Nesse caso, curvas de regressão são ajustadas para cada indivíduo e são admitidos coeficientes de regressão variando aleatoriamente entre os indivíduos. Essa variação ocorre, de forma simples, nos interceptos ou, de forma mais complexa, também nas inclinações. O modelo de coeficientes aleatório é um caso especial do modelo misto, pois os coeficientes de regressão das curvas individuais serão considerados efeitos aleatórios e os coeficientes de regressão das curvas médias, efeitos fixos (Hauser et al., 2009).

Assim, para se estruturar um modelo misto é necessário ter um modelo para os efeitos fixos, que deve descrever o comportamento da curva média; um modelo para os efeitos aleatórios, que junto com os efeitos fixos deve descrever as curvas individuais e um modelo para dispersão condicional, que deve descrever as variâncias-covariância das observações individuais ao longo do tempo (Suyama, 2005).

2.2.4.1. Especificação do modelo misto

O modelo linear misto é denotado pela seguinte forma matricial:

$$Y = X\beta + Zb + \varepsilon$$

em que:

$Y =$ é o vetor de observações ou vetor ($n \times 1$) de respostas da i -ésima unidade experimental ou indivíduo;

$X =$ é uma matriz ($n \times p$) conhecida e de especificação, de posto completo dos efeitos fixos;

$\beta =$ é um vetor ($p \times 1$) de parâmetros de efeitos fixos desconhecidos;

$Z =$ é uma matriz ($n \times q$) conhecida e de especificação, de posto completo dos efeitos aleatórios;

$b =$ é um vetor ($q \times 1$) de parâmetros de efeitos aleatórios desconhecidos;

$\varepsilon =$ é o vetor ($n \times 1$) de erros aleatórios não observáveis.

As matrizes de especificação X e Z podem ser diferentes e variar entre unidades experimentais, estendendo o modelo para o caso de dados não balanceados em relação ao tempo (Barbosa, 2009). A matriz Z pode conter quaisquer covariáveis que afetem diferentemente as unidades experimentais (Lima, 1996). Segundo Amador (2010) as colunas das matrizes X e Z podem estar associadas:

- aos fatores que definem a estrutura das subpopulações (grupos ou tratamentos);
- ao fator tempo, identificando, por exemplo, a forma da curva a ser ajustada;
- a covariáveis, cujos efeitos na resposta média desejam pesquisar.

Assume-se que os efeitos aleatórios (b) e os erros (resíduos) (ε) são não correlacionados, têm distribuição normal, média zero ($E(b) = 0$ e $E(\varepsilon) = 0$) e matrizes de variâncias e covariâncias, respectivamente, G e R , ou seja, $b \sim N(0, G)$ e $\varepsilon \sim N(0, R)$. As matrizes G e R são matrizes positivas definidas, por hipótese, e, portanto, não singulares, dadas por:

$$\text{Var}(b) = E(bb') = G \text{ e } \text{Var}(\varepsilon) = E(\varepsilon\varepsilon') = R$$

As esperanças e variâncias são dadas pelas matrizes:

$$E \begin{bmatrix} b \\ \varepsilon \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{e} \quad \text{Var} \begin{bmatrix} b \\ \varepsilon \end{bmatrix} = \begin{bmatrix} G & 0 \\ 0 & R \end{bmatrix}$$

Deste modo tem-se que:

$$E(\mathbf{Y}) = E(\mathbf{X}\beta + \mathbf{Z}b + \varepsilon) = \mathbf{X}\beta \text{ e};$$

$$\text{Var}(\mathbf{Y}) = \text{Var}(\mathbf{X}\beta + \mathbf{Z}b + \varepsilon) = \text{Var}(\mathbf{X}\beta) + \text{Var}(\mathbf{Z}b) + \text{Var}(\varepsilon) = \mathbf{Z}\text{Var}(b)\mathbf{Z}' + \mathbf{R} =$$

$$\text{Var}(\mathbf{Y}) = \mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R}$$

$$\text{Assim, } \mathbf{Y} \sim N(\mathbf{X}\beta; \mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R})$$

2.2.4.2. Estimação de β e Predição de b

Nos modelos mistos têm-se, em geral, interesse na estimação dos efeitos fixos, predição de efeitos aleatórios e estimação dos componentes de variância e covariância. Assim, em um modelo misto a estimação dos efeitos fixos (β) e predição dos efeitos aleatórios (b), podem ser obtidas utilizando-se o Método de Estimação dos Mínimos Quadrados Generalizados (GLSE: Generalized Least Square Estimation) que conduz ao sistema de equações normais do modelo misto (Amador, 2010).

O Método de Estimação dos Mínimos Quadrados Generalizados (GLSE: Generalized Least Square Estimation) ou Equações de Estimação Generalizadas (Generalized Estimating Equations – GEE) foi proposto por Liang e Zeger (1986) com o objetivo de estimar parâmetros de regressão especialmente quando os dados estão correlacionados. Este método ao considerar a estrutura de correlação entre as observações produz estimativas mais eficientes e não viciadas para os parâmetros do modelo de regressão (Agranonik, 2009).

Para estimar parâmetros e prever efeitos aleatórios, a estrutura especial da distribuição normal para b e ε permite dividir o desenvolvimento teórico em duas etapas. Primeiro considera-se G e R conhecidas, obtendo-se solução para β e b . Depois, estimam-se os parâmetros não conhecidos em G e R . Computacionalmente, a segunda etapa é executada primeiro (Riboldi, 2007).

A estimação dos efeitos fixos e a predição dos efeitos aleatórios são realizadas através da derivação das equações do modelo misto (EMM). Essa derivação pode ser feita pela minimização da soma de quadrados dos resíduos ou pela maximização da função densidade de probabilidade conjunta de Y e b , sendo a última mais adotada:

$$f(y) = \frac{1}{(2\pi)^{n/2} (ZGZ' + R)^{1/2}} e^{-\frac{1}{2}[(y-X\beta)'(ZGZ'+R)^{-1}(y-X\beta)]}$$

A função densidade de probabilidade conjunta de Y e b pode ser escrita como o produto entre a função densidade condicional de Y , dado b , e a função densidade de probabilidade de b . Desta forma tem:

$$f(Y, b) = f(Y/b) \cdot f(b)$$

Assim:

$$f(Y, b) = \frac{1}{(2\pi)^{n/2} |R|^{1/2}} e^{-\frac{1}{2}[(y-X\beta-Zb)'(R)^{-1}(y-X\beta-Zb)]} \cdot \frac{1}{(2\pi)^{g/2} |G|^{1/2}} e^{-\frac{1}{2}[(b-0)'(G)^{-1}(b-0)]}$$

Para se proceder à maximização de $f(Y, b)$, pode-se usar a transformação por logaritmo. Assim, fazendo-se $L = \log [f(Y, b)]$, tem-se:

$$L = \frac{1}{2} 2n \log(2\pi) - \frac{1}{2} (\log R + \log G) - \frac{1}{2} (y' R^{-1} y - 2 y' R^{-1} X\beta - 2 y' R^{-1} Zb + 2 \beta' X' R^{-1} Zb + \beta' X' R^{-1} X\beta + b' Z' R^{-1} Zb + b' G^{-1} b)$$

As equações do modelo misto (EMM), com as quais são fornecidas as soluções para os efeitos fixos e predição para os efeitos aleatórios são obtidas derivando-se L em relação a β e b, e tomando-se tais derivadas identicamente nulas:

$$\begin{bmatrix} \frac{\partial L}{\partial \beta} \\ \frac{\partial L}{\partial b} \end{bmatrix} = \begin{bmatrix} -X'R^{-1}y + X'R^{-1}X\beta^\circ + X'R^{-1}Z\hat{b} \\ -Z'R^{-1}y + Z'R^{-1}X\beta^\circ + Z'R^{-1}Z\hat{b} + G^{-1}\hat{b} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} X'R^{-1}X\beta^\circ + X'R^{-1}Z\hat{b} \\ -Z'R^{-1}X\beta^\circ + Z'R^{-1}Z\hat{b} + G^{-1}\hat{b} \end{bmatrix} = \begin{bmatrix} -X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix}$$

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ -Z'R^{-1}X & Z'R^{-1}Z + G^{-1} \end{bmatrix} \begin{bmatrix} \beta^\circ \\ \hat{b} \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix}$$

As soluções são dadas por:

$$\begin{bmatrix} \beta^\circ \\ \hat{b} \end{bmatrix} = \begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ -Z'R^{-1}X & Z'R^{-1}Z + G^{-1} \end{bmatrix}^{-1} \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix}$$

Quando G e R são conhecidas o estimador de β e o preditor de b, são, respectivamente, o estimador de Mínimos Quadrados Generalizados (GLSE), o BLUE e o melhor preditor linear imparcial (BLUP). No entanto, como já mencionado, G e R geralmente são desconhecidas, dispondo-se apenas de estimativas obtidas por algum método. Neste caso, os termos BLUE E BLUP não mais se aplicam, sendo apropriado substituí-los por EBLUE (empirical best linear unbiased estimator) e EBLUP (empirical best linear unbiased estimator predictor) (Littell et al., 1996).

2.2.4.3. Estimação dos parâmetros de covariância em G e R

O emprego das equações do Modelo Misto (EMM) depende do conhecimento da matriz V definida em função de G e R . Conforme visto anteriormente a matriz $V = ZGZ' + R$ está associada à matriz Z (matriz conhecida de efeitos aleatórios) e às matrizes G e R , matriz de covariância de efeitos aleatórios e matriz de covariância do erro, respectivamente. Apesar do conhecimento de diversos tipos de estruturas de matriz de variância e covariâncias, que podem ser escolhidas de forma a representar a variabilidade dos dados da forma mais real possível, seus componentes não são conhecidos, sendo necessário substituí-los por suas estimativas.

Existem vários métodos de estimação de componentes de variâncias, dentre os quais pode destacar: o método da Análise da Variância (Fischer, 1918), os Métodos I, II e III de Henderson (Henderson, 1953), Método de Estimação Quadrática Não-Viesada de Norma Mínima - MINQUEO (Rao, 1970; Rao, 1971), Método de Estimação Quadrática Não-Viesada de Variância Mínima – MIVQUE (Rao, 1971b), Método da Máxima Verossimilhança (ML) (Hartley e Rao, 1967) e Método da Máxima Verossimilhança Restrita (REML) (Patterson e Thompson, 1971). No entanto, segundo Pinheiro e Bates (2000) os métodos ML e REML são os mais comumente usados para se estimar os componentes de variância.

Em várias situações, a preferência tem sido dada aos métodos baseados em verossimilhança, os quais exploram a suposição de que b e ε têm distribuição normal (Verneque, 1994; Littell et al., 1996). Os estimadores de máxima verossimilhança baseiam-se na obtenção do ponto de máximo da função de verossimilhança (função de densidade de probabilidade), por meio da derivação dessa função em relação ao parâmetro de interesse e apresentam as seguintes propriedades desejáveis: translação invariante (não afetados por mudanças nos efeitos fixos), suficiência, consistência e eficiência. Além disso, possibilitam a geração de estimativas não negativas de componentes de variância.

O método da máxima verossimilhança (ML) foi desenvolvido por Fisher segundo Searle (1987), e primeiramente usado para a obtenção de estimativas de componentes de variância, em modelos mistos, em Hartley e Rao (1967). O método utiliza a distribuição dos dados observados, a qual aplica uma função de verossimilhança que, por sua vez, é função dos parâmetros do modelo. Os

estimadores de máxima verossimilhança são os valores dos parâmetros que maximizam a função de verossimilhança (Garcia, 2004). Apesar de apresentar diversas propriedades desejáveis os estimadores ML são viciados em decorrência da imposição de restrição de não negatividade e perda de graus de liberdade devido à estimação dos efeitos fixos (Shaw, 1987).

Visando eliminar a fonte de vício, relacionada à perda de graus de liberdade, Patterson e Thompson (1971) apresentaram uma modificação ao método ML, o qual passou a ser denominado de máxima verossimilhança restrita ou residual (REML), em geral, o método mais importante para estimar parâmetros de covariância (Amador, 2010). Além de apresentar as propriedades desejáveis do método da ML e também a imposição da restrição de não negatividade, os estimadores REML não são viciados. Isto ocorre porque segundo Amador (2010), a maximização da função densidade de probabilidade da parte referente aos efeitos aleatórios, em relação aos componentes de variância, elimina o viés resultante da perda de graus de liberdade na estimação dos efeitos fixos do modelo.

No REML, cada observação é dividida em duas partes independentes, uma referente aos efeitos fixos e outra aos aleatórios, de maneira que a função de densidade de probabilidade das observações é dada pela soma das funções de densidade de probabilidade de cada parte (Riboldi, 2007). Os parâmetros do modelo são obtidos pela maximização da função de verossimilhança das duas partes, ou seja, a referente aos efeitos fixos e aos efeitos aleatórios. Maximizando a função de verossimilhança da parte referente aos efeitos aleatórios obtém os estimadores para os componentes de variância, enquanto que a maximização da parte referente aos efeitos fixos fornece os estimadores para os efeitos fixos (Garcia, 2004). Assim, o procedimento REML sob modelo misto realiza, simultaneamente, as operações de estimação (BLUE) dos efeitos fixos por quadrados mínimos generalizados, de estimação de componentes de variância por REML e de predição (BLUP) de efeitos aleatórios (Garcia, 2004).

Freitas (2007) e Macedo (2007) apresenta o log da função de verossimilhança para o método da máxima verossimilhança restrita (REML).

$$l_r(G, R) = -\frac{1}{2} \log|V| - \frac{1}{2} \left| \log X'V^{-1}X \right| - \frac{n-p}{2} \log(y - X(X'V^{-1}X)^{-1}X'V^{-1}y)'V^{-1} \\ (y - X(X'V^{-1}X)^{-1}X'V^{-1}y) - \frac{n-p}{2} \left[1 + \log\left(\frac{2\pi}{n-p}\right) \right]$$

Para maximizar o logaritmo da verossimilhança restrita e, desta forma, estimar os componentes de variância G e R são necessários métodos iterativos, como o método Newton-Raphson, o método Fisher scoring e o método EM proposto por Laird e Ware (1982). Porém, segundo Barbosa (2009) o método Newton-Raphson com as modificações propostas por Jennrich e Schluchter (1986), é considerado melhor que os demais em relação ao tempo total para a convergência.

2.2.4.4. Estruturas de Covariância

A escolha da melhor estrutura da matriz de covariância visa obter uma estrutura parcimoniosa, que explique bem a variabilidade dos dados das diversas ocasiões e a correlação entre essas medidas com um número pequeno de parâmetros, o que pode melhorar a eficiência das inferências sobre os parâmetros do modelo proposto para os valores médios nas diversas ocasiões (Barbosa, 2009). A escolha adequada da estrutura de covariância é necessária, pois afeta as estimativas e os erros padrões de efeitos fixos, diagnósticos, interpretações e extrapolações. A escolha depende de informação empírica, da estrutura dos dados e de programas computacionais disponíveis com a estrutura selecionada (Castro, 1999). A seguir serão apresentadas algumas das estruturas de covariância, para as matrizes G e R, mais utilizadas e implementadas em software estatísticos como o SAS e o R, todas elas com t=4 ocasiões.

Componente de Variância (VC)

Variâncias iguais em todas as ocasiões de medidas e observações independentes e envolve um único parâmetro.

$$\begin{pmatrix} \sigma^2 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & \sigma^2 \end{pmatrix}$$

Simetria Composta (CS)

Variâncias iguais em todas as ocasiões de medidas e mesmas covariâncias entre medidas feitas em ocasiões distintas e envolve dois parâmetros.

$$\begin{pmatrix} (\sigma^2 + \sigma_1^2) & \sigma_1^2 & \sigma_1^2 & \sigma_1^2 \\ \sigma_1^2 & (\sigma^2 + \sigma_1^2) & \sigma_1^2 & \sigma_1^2 \\ \sigma_1^2 & \sigma_1^2 & (\sigma^2 + \sigma_1^2) & \sigma_1^2 \\ \sigma_1^2 & \sigma_1^2 & \sigma_1^2 & (\sigma^2 + \sigma_1^2) \end{pmatrix}$$

Sem Estrutura ou não estruturada (UN)

Todas as variâncias e as covariâncias podem ser desiguais. Especifica uma matriz completamente geral, parametrizada diretamente em termos de variâncias e covariâncias. As variâncias são restritas a valores não negativos e as covariâncias não têm restrições. O número de parâmetros é dado por $t(t+1) / 2$.

$$\begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 & \sigma_{34} \\ \sigma_{14} & \sigma_{24} & \sigma_{34} & \sigma_4^2 \end{pmatrix}$$

Diagonal Principal “*Banded*” (UN (1))

Variâncias desiguais e covariâncias nulas. O número de parâmetros é igual a t .

$$\begin{pmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{pmatrix}$$

Autorregressiva de Primeira Ordem (AR (1))

Essa estrutura indica que todos os níveis de tempo têm a mesma variância. Indica também, que a correlação entre as medidas tomadas no tempo vai decrescendo à medida que aumenta a distância entre os níveis de tempo. Isto quer dizer que as medidas tomadas em tempos mais distantes têm menos correlação que as medidas tomadas em tempos mais próximos. Na estrutura AR (1) o número de parâmetros envolvidos é igual a 2.

$$\sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{pmatrix}$$

Autorregressiva de Primeira Ordem Heterogênea (ARH (1))

É uma generalização da estrutura AR (1), impondo variâncias e covariâncias desiguais. A estrutura ARH (1) somente é viável se os dados são igualmente espaçados no tempo e envolve $t + 1$ parâmetro.

$$\begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho^2 & \sigma_1\sigma_4\rho^3 \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_2\sigma_3\rho & \sigma_2\sigma_4\rho^2 \\ \sigma_3\sigma_1\rho^2 & \sigma_3\sigma_2\rho & \sigma_3^2 & \sigma_3\sigma_4\rho \\ \sigma_4\sigma_1\rho^3 & \sigma_4\sigma_2\rho^2 & \sigma_4\sigma_3\rho & \sigma_4^2 \end{pmatrix}$$

Simetria Composta Heterogênea (CSH)

Variâncias diferentes para cada elemento da diagonal principal e, a raiz quadrada desses parâmetros nos elementos fora da diagonal principal, sendo: σ_i^2 , o i-ésimo parâmetro de variância e ρ o parâmetro de correlação satisfazendo $|\rho| < 1$. Envolve $t + 1$ parâmetro.

$$\begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho & \sigma_1\sigma_4\rho \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_2\sigma_3\rho & \sigma_2\sigma_4\rho \\ \sigma_3\sigma_1\rho & \sigma_3\sigma_2\rho & \sigma_3^2 & \sigma_3\sigma_4\rho \\ \sigma_4\sigma_1\rho & \sigma_4\sigma_2\rho & \sigma_4\sigma_3\rho & \sigma_4^2 \end{pmatrix}$$

Toeplitz (TOEP)

Estrutura de covariância de um processo estocástico de médias móveis de ordem $q = (t - 1)$, neste exemplo $q = 3$.

$$\begin{pmatrix} \sigma^2 & \sigma_1 & \sigma_2 & \sigma_3 \\ \sigma_1 & \sigma^2 & \sigma_1 & \sigma_2 \\ \sigma_2 & \sigma_1 & \sigma^2 & \sigma_1 \\ \sigma_3 & \sigma_2 & \sigma_1 & \sigma^2 \end{pmatrix}$$

Antidependência de Ordem 1 (ANTE (1))

Impõe parâmetros de variância diferentes para cada elemento da diagonal, sendo os elementos fora da diagonal principal, funções de variâncias e do K-ésimo parâmetro de autocorrelação, satisfazendo ($|\rho_k| < 1$). O número de parâmetros é dado por $2t-1$.

$$\begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_1 & \sigma_1\sigma_3\rho_1\rho_2 & \sigma_1\sigma_4\rho_1\rho_2\rho_3 \\ \sigma_2\sigma_1\rho_1 & \sigma_2^2 & \sigma_2\sigma_3\rho_2 & \sigma_2\sigma_4\rho_2\rho_3 \\ \sigma_3\sigma_1\rho_2\rho_1 & \sigma_3\sigma_2\rho_2 & \sigma_3^2 & \sigma_3\sigma_4\rho_3 \\ \sigma_4\sigma_1\rho_3\rho_2\rho_1 & \sigma_4\sigma_3\rho_3\rho_2 & \sigma_4\sigma_3\rho_3 & \sigma_4^2 \end{pmatrix}$$

Esta estrutura permite que as variâncias sejam diferentes e é aplicável em estudos longitudinais em que as condições de avaliação não são igualmente espaçadas, e apresenta heterogeneidade de variância e correlação serial (Barbosa, 2009).

Correlação sem Estrutura (UNR)

Especifica uma matriz de variância-covariância completamente geral em termos de variâncias e correlações. Essa estrutura ajusta o mesmo modelo que o tipo UM, mas com diferente parametrização. Sendo σ_i^2 o i-ésimo parâmetro de variância e ρ_{jk} é a correlação entre a j-ésima e a k-ésima medida satisfazendo $|\rho_{jk}| < 1$. O número de parâmetros é dado por $t(t+1)/2$.

$$\begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{21} & \sigma_1\sigma_3\rho_{31} & \sigma_1\sigma_4\rho_{41} \\ \sigma_2\sigma_1\rho_{21} & \sigma_2^2 & \sigma_2\sigma_3\rho_{31} & \sigma_2\sigma_4\rho_{42} \\ \sigma_3\sigma_1\rho_{31} & \sigma_3\sigma_2\rho_{32} & \sigma_3^2 & \sigma_3\sigma_4\rho_{43} \\ \sigma_4\sigma_1\rho_{41} & \sigma_4\sigma_2\rho_{42} & \sigma_4\sigma_3\rho_{43} & \sigma_4^2 \end{pmatrix}$$

Huynh-Feldt (H-F)

É similar à estrutura CSH, pois tem a mesma heterogeneidade não estruturada ao longo da diagonal principal e o mesmo número de parâmetros desconhecidos, entretanto, a estrutura HF constrói os elementos fora da diagonal, calculando uma média aritmética. O número de parâmetros é dado por $t+1$.

$$\begin{bmatrix} \sigma_1^2 & \frac{(\sigma_1^2 + \sigma_2^2)}{2} - \lambda & \frac{(\sigma_1^2 + \sigma_3^2)}{2} - \lambda & \frac{(\sigma_1^2 + \sigma_4^2)}{2} - \lambda \\ \frac{(\sigma_1^2 + \sigma_2^2)}{2} - \lambda & \sigma_2^2 & \frac{(\sigma_2^2 + \sigma_3^2)}{2} - \lambda & \frac{(\sigma_2^2 + \sigma_4^2)}{2} - \lambda \\ \frac{(\sigma_3^2 + \sigma_1^2)}{2} - \lambda & \frac{(\sigma_3^2 + \sigma_2^2)}{2} - \lambda & \sigma_3^2 & \frac{(\sigma_3^2 + \sigma_4^2)}{2} - \lambda \\ \frac{(\sigma_4^2 + \sigma_1^2)}{2} - \lambda & \frac{(\sigma_4^2 + \sigma_2^2)}{2} - \lambda & \frac{(\sigma_4^2 + \sigma_3^2)}{2} - \lambda & \sigma_4^2 \end{bmatrix}$$

Toeplitz banded (TOEP2)

$$\begin{pmatrix} \sigma^2 & \sigma_1 & 0 & 0 \\ \sigma_1 & \sigma^2 & \sigma_1 & 0 \\ 0 & \sigma_1 & \sigma^2 & \sigma_1 \\ 0 & 0 & \sigma_1 & \sigma^2 \end{pmatrix}$$

Além dessas estruturas de variância-covariância, outras podem ser encontradas nos softwares estatísticos SAS e R.

2.2.5. Seleção de Modelos

Selecionar o melhor modelo significa, não só selecionar a melhor estrutura para as médias (parte fixa), como também a melhor estrutura de covariância. Para esse propósito Rocha (2004) propõe uma série de técnicas gráficas e analíticas que auxiliam na escolha das matrizes dos modelos lineares mistos, afirmando que para estudos longitudinais, é razoável utilizar informações sobre o comportamento

da resposta ao longo das ocasiões de avaliação na modelagem da estrutura de covariância intra-unidades amostrais. No entanto, um critério objetivo de seleção deve ser seguido.

Muitos métodos já foram desenvolvidos visando facilitar a escolha da estrutura de covariância que melhor explique o comportamento da variabilidade e da correlação entre as medidas repetidas e, dessa forma, selecionar o modelo adequado (Floriano et al., 2006). Destacam-se os Critérios de Informação de Akaike – AIC e Bayesiano – BIC. Além desses, outros dois testes que podem ser utilizados são a estatística Z de Wald e o Teste Assintótico da Razão de Verossimilhanças (LR). Porém, segundo Camarinha Filho (2002) a estatística Z de Wald pode ser incerta para pequenos conjuntos de dados e para parâmetros que apresentam distribuição assimétrica ou distribuição amostral limite. Quanto ao LR a desvantagem é que só pode ser usado para comparar dois modelos de cada vez, sendo que um é caso especial do outro (Xavier, 2000, Camarinha Filho, 2002).

Os critérios de AIC e BIC são, na verdade, valores para os logaritmos das funções de verossimilhanças do modelo, que dependem do número de observações e do número de parâmetros do modelo (Costa, 2003). O critério de AIC é calculado por $AIC = -2L(\hat{\theta}) + 2d$; o critério de BIC é calculado por $BIC = -2L(\hat{\theta}) + \ln(N)d$ e o critério de Akaike corrigido (AICc) é dado por $AICc = -2L(\hat{\theta}) + 2d + 2d(d+1)(n-d-1)^{-1}$ em que: $L(\hat{\theta})$ é o ln da função de verossimilhança, d representa o número total de parâmetros de efeito fixo e aleatório estimado do modelo e $N = \sum n_i$ é o somatório de todas as observações. Em ambos os critérios o melhor modelo será o que apresentar o menor valor para o(s) critério utilizado.

Quanto à utilização dos critérios de AIC e BIC, West et al. (2007) alertam para o fato de que alguns softwares calculam os valores de AIC e BIC utilizando equações diferentes, dependendo se a estimação é feita por ML ou REML. Além disso, Bates e Pinheiro (2000) discutem que quando os modelos são ajustados por REML, os valores de AIC e BIC somente podem ser comparados entre modelos com mesma estrutura de efeitos fixos. Já quando os modelos são ajustados por ML, os valores de AIC e BIC podem ser comparados com quaisquer modelos ajustados para os mesmos dados.

2.2.6. Teste Assintótico da Razão de Verossimilhanças (LR)

A estatística $-2\log Veross$ é baseada no logaritmo da razão entre as duas verossimilhanças: do modelo mais simples, $\iota(\tilde{\theta})$, e o modelo mais complexo, $\iota(\hat{\theta})$.

$$-2\log Veross = -2 [\log(\iota(\tilde{\theta})) - \log \iota(\hat{\theta})] \sim \chi_r^2$$

É assintoticamente distribuída como uma qui-quadrado com r graus de liberdade e serve para testar a hipótese H_0 : o modelo mais simples é adequado. Quando esta hipótese for rejeitada, conclui-se que o modelo mais complexo (com maior número de parâmetros) é adequado. Se o modelo mais simples for adequado, os valores da função de verossimilhança avaliada em $\tilde{\theta}$ e $\hat{\theta}$ devem estar próximos, indicando que os dados estão dando suporte ao modelo com menor número de parâmetros. Embora seja bastante eficiente, a principal desvantagem desse teste é que ele só pode ser usado para comparar dois modelos de cada vez, sendo que um deles deve ser um caso especial do outro.

3.0. MATERIAL E MÉTODOS

3.1. Programa de melhoramento do feijão vagem da Universidade Estadual do Norte Fluminense Darcy Ribeiro (UENF)

O programa de melhoramento de feijão vagem da Universidade Estadual do Norte Fluminense Darcy Ribeiro (UENF), cujo objetivo é a seleção de genótipos produtivos e de qualidade comercial para as regiões Norte e Noroeste Fluminense se iniciou com a caracterização e o estudo da diversidade genética de 25 acessos de feijão vagem de hábito de crescimento indeterminado do Banco de Germoplasma da UENF. A partir de então foram realizados os cruzamentos entre cinco acessos divergentes e com características desejáveis, obtendo dez híbridos dialélicos. Foram realizadas seleções nas populações F2, em campo; avançando as gerações F3, F4 e F5 pelo método SSD (“single seed descent” - descendente de uma única semente por planta), em casa de vegetação, abrindo e selecionando linhagens em F6. A partir daí foram selecionadas 27 linhagens promissoras desta geração F6 com a qual se realizou um trabalho em três estações experimentais: Campos dos Goytacazes (RJ); Itaocara (RJ) e Bom Jesus do Itabapoana (RJ) obtendo a geração F7. Em sequência buscou-se a geração F8 realizando um novo experimento utilizando 27 linhagens selecionadas de feijão vagem da geração F7 e mais três testemunhas (duas variedades comerciais, Feltrin, Top Seed Blue Line e um dos progenitores, 19 UENF-1445), de hábito de crescimento indeterminado, do programa de melhoramento da UENF (Araujo, 2011).

covariâncias para G e R. Primeiramente fixou-se para R a estrutura do tipo componente de variâncias (VC) variando-se as demais estruturas definidas. Posteriormente fixou-se a estrutura selecionada para R e variou-se as demais visando escolher a mais adequada para G. Cada combinação de estrutura de covariância (G-R) compreendeu um modelo. Desta forma, para cada grupo foram formados 18 modelos, sendo todos com a mesma estrutura fixa e aleatória diferindo somente nas estruturas de covariâncias consideradas.

O método utilizado para estimação dos componentes de variâncias foi o da máxima verossimilhança restrita (REML), ao passo que a estatística F foi utilizada para testar a significância dos efeitos fixos do modelo.

Os modelos foram gerados mediante a utilização de um polinômio de segundo grau e selecionados pelos critérios de informação de Akaike (AIC), Akaike corrigido (AICC), Bayesiano (BIC) e pelo valor que maximiza o logaritmo da função de verossimilhança (-2Res log Likelihood). De forma a auxiliar no processo de escolha do melhor modelo dentro de cada grupo foi utilizado o teste assintótico da razão de verossimilhança, de acordo com Littell et al. (2006), para selecionar, entre duas estruturas selecionadas, a mais adequada. As análises estatísticas foram realizadas com o software SAS 9.1.3 e os gráficos com o software R 3.0.0.

Dos experimentos realizados, foram selecionadas 14 linhagens de feijão vagem, da geração F8. Dando segmento ao programa de melhoramento genético da UENF, foi realizado um trabalho com as linhagens selecionadas e três cultivares de feijão vagem (duas variedades comerciais e um dos progenitores da UENF), em Bom Jesus do Itabapoana, RJ, para obtenção da geração F9 (Araujo, 2011; Almeida, 2011).

3.2 Material

Com a finalidade de se aplicar a metodologia de modelos lineares mistos para dados longitudinais utilizou-se um conjunto de dados proveniente de um experimento realizado no Instituto Federal Fluminense, Coordenadas UTM 224021m E e 7660399m S, zona 24K, altitude 84m, com temperatura média anual de 23°C e precipitação média anual em torno de 1480mm (IBGE, 2012), localizado no campus de Bom Jesus do Itabapoana – RJ, Campos dos Goytacazes – RJ.

O experimento consistiu na avaliação de 17 genótipos de feijão vagem (*Phaseolus vulgaris* L.) de hábito de crescimento indeterminado, sendo 2 variedades comerciais e 15 genótipos do programa de melhoramento de feijão vagem da UENF. Dos 17 genótipos, 3 são testemunhas e 14 são linhagens selecionadas a partir da geração F8 do programa de melhoramento de feijão vagem da UENF. As 3 testemunhas foram formadas pelo Progenitor UENF 19 (UENF 1445) e as cultivares comerciais Feltrin e Top Seed Blue Line. As linhagens foram: UENF 7-3-1, UENF 7-4-1, UENF 7-5-1, UENF 7-6-1, UENF 7-9-1, UENF 7-10-1, UENF 7-10-1, UENF 7-10-1, UENF 7-12-1, UENF 7-14-1, UENF 7-20-1, UENF 9-24-2, UENF 14-3-3, UENF 14-4-3, UENF 14-6-3 e UENF 15-23-4.

O experimento foi implantado em um delineamento de blocos ao acaso com quatro repetições e 10 plantas por parcela no espaçamento de 1,0 x 0,5 m considerando o esquema de parcela subdividida no tempo.

O plantio foi feito em área que já havia sido cultivada com feijão vagem. O preparo do solo foi realizado pelo método convencional, com uma aração e duas gradagens. As características químicas do solo reveladas pela análise de solo foram: pH = 5,7; fósforo = 14 mg.dm⁻³; potássio = 211 mg.dm⁻³; enxofre = 4,0

mg.dm⁻³; Ca = 4,3 cmol.dm⁻³; Mg = 1,5 cmol.dm⁻³; Al = 0,0 cmol.dm⁻³; H+Al = 3,7 g.kg⁻¹; C = 11,5 g.kg⁻¹e M.O = 19,8 g.kg⁻¹.

Considerando as características químicas do solo e a necessidade da cultura de acordo com Carrijo et al. (1999b) aplicou-se 150 kg ha⁻¹ de N, sendo 30% na semeadura e o restante em cobertura (30 e 60 dias após o plantio), usando ureia. A adubação potássica foi realizada mediante a aplicação de 30 kg ha⁻¹ de K₂O (cloreto de potássio), sendo 50% no plantio e o restante juntamente com as demais aplicações de nitrogênio. O fósforo foi aplicado todo no plantio, na dose de 180 kg ha⁻¹ de P₂O₅ (superfosfato simples).

A semeadura foi realizada em 12 de junho de 2012, colocando-se três sementes por cova, a uma profundidade de 2,5/cm. A emergência iniciou, em média, dez dias após a semeadura, quando foi realizado o desbaste deixando-se apenas uma planta por cova. Cerca de quinze dias após a emergência, as plantas foram tutoradas com bambu e arame. Durante a condução do experimento, foram realizados os tratos culturais e fitossanitários recomendados para a cultura, segundo Filgueira (2000), bem como irrigação por aspersão.

A variável utilizada foi a Altura de Plantas, mensurada em centímetros (cm) em intervalos de quatro dias com régua de madeira de dois metros. As avaliações se iniciaram 12 dias após o plantio e foi até as plantas atingirem a altura máxima do espaldeamento, o que ocorreu 32 dias após a primeira avaliação para os genótipos de maiores taxas de crescimento.

As avaliações realizadas caracterizaram as medidas no tempo, sendo, portanto 9 medidas no tempo. Foram consideradas as medidas tomadas nas 5 plantas centrais de cada linha. Como as medidas foram feitas sempre sobre a mesma unidade experimental, o esquema ficou caracterizado como “parcelas subdivididas no tempo” ou “ensaio de medidas repetidas”.

3.3. Métodos

Foram consideradas diversas abordagens estatísticas, visando escolher dentre elas a mais adequada para o estudo dos dados longitudinais. As abordagens estatísticas foram: análise univariada de perfil (parcelas subdivididas no tempo), análise multivariada e modelos mistos.

3.3.1. Análise Univariada de Perfil

Nessa abordagem os dados foram analisados sob o enfoque das parcelas subdivididas, onde o fator da parcela (entre indivíduos) é aquele cujos níveis permanecem constantes, enquanto o fator da subparcela (intra-indivíduos) entre eles varia. A parcela foi representada pelos genótipos e blocos, ao passo que a subparcela foi formada pelo fator tempo (dias) e a interação tempo x genótipos.

As hipóteses testadas foram a de perfis coincidentes (hipótese de diferença entre genótipos), perfis paralelos (hipótese da interação entre genótipo e tempo) e perfis horizontais (hipótese do efeito do tempo):

O modelo adotado foi:

$$Y_{ijk} = \mu + \beta_j + \tau_i + (\beta\tau)_{ji} + \gamma_k + (\tau\gamma)_{ik} + \varepsilon_{ijk},$$

em que:

Y_{ijk} : é o valor observado para a variável Altura de plantas no **k-ésimo** tempo (dias) para o **i-ésimo** genótipo no **j-ésimo** bloco;

μ : é uma constante inerente a todas as observações;

β_j : é o efeito do bloco i , $i = 1, 2, 3$ e 4 ;

τ_i : é o efeito do genótipo j , $j = 1, 2, 3, \dots, 17$;

$(\beta\tau)_{ji}$: é o efeito aleatório da interação do i -ésimo bloco com o j -ésimo genótipo (erro associado às parcelas);

γ_k : é o efeito do tempo k , $k = 1, 2, 3, \dots, 9$;

$(\tau\gamma)_{ik}$: é o efeito da interação entre o j -ésimo genótipo com o k -ésimo tempo;

ε_{ijk} : é o erro aleatório correspondente às observações do k -ésimo tempo para o i -ésimo bloco no j -ésimo genótipo (variação do acaso sobre as observações), supostos homocedásticos, independentes e normalmente distribuídos.

Para que haja validação dos resultados, uma pressuposição dessa técnica é que a matriz de variâncias e covariâncias (Σ) atenda à condição de Huynh e Feldt (1970). Para verificar essa pressuposição, foi utilizado o teste de esfericidade de Mauchly. Se a matriz de covariâncias não atender à condição de Huynh e Feldt (1970), pelo teste Mauchly, a primeira alternativa considerada é realizar o ajuste dos graus de liberdade do teste F das fontes de variação intra-indivíduo utilizando as correções de Geisser e Greenhouse (1958) e também Huynh e Feldt (1976), realizando, portanto uma análise univariada aproximada com correção dos graus de liberdade. A segunda alternativa da análise foi a realização da análise multivariada de perfil.

3.2. Análise multivariada de perfil

A análise de variância multivariada tem sido aplicada na análise de experimentos com medidas repetidas para contemplar a covariância entre observações ao longo do tempo e na mesma parcela. Nesta análise, as medidas em cada nível do fator tempo (cortes) são consideradas como variáveis distintas, explorando-se as variâncias e covariâncias observadas entre elas. Assim, não se exige qualquer suposição sobre a estrutura de Σ , o que a caracteriza como uma matriz não estruturada ($\Sigma = UN$) (Silva, 2011). Nesse sentido, o seguinte modelo matemático, sugerido por Vonesh & Chinchilli (1997), foi analisado (os termos deste modelo já foram definidos no modelo univariado):

$$Y_{ijk} = \mu + \beta_i + \tau_j + \gamma_k + (\tau\gamma)_{jk} + \varepsilon_{ijk},$$

As estatísticas de teste utilizadas na análise multivariada foram obtidas através dos princípios da união-intersecção de Roy e da razão de verossimilhança de Wilks. Definindo-se $\theta_i = \lambda_i (1 + \lambda_i)^{-1}$, em que λ_i é a i -ésima raiz característica (autovalor) de \mathbf{HE}^{-1} , sendo: \mathbf{H} matriz da soma de quadrados e produtos cruzados devido à hipótese nula; e \mathbf{E} a matriz de soma de quadrados e produtos cruzados devido ao erro (Xavier, 2000). As estatísticas de teste consideradas foram:

- i) Pillai's Trace: $P = \sum \theta_i$, com $i = 1, 2, \dots, s$; sendo $s = \min(t-1, g-1)$; em que, 't' é o número de tempos; e 'g', o de tratamentos;
- ii) Wilks' Lambda: $\Delta = \prod (1 - \theta_i)$;
- iii) Hotelling – Lawley Trace: $T = \sum \theta_i (1 - \theta_i)^{-1}$;
- iv) Roy's Greatest Root: $\theta_s = \max(\theta_i)$.

As distribuições exatas dessas estatísticas, sob a hipótese nula, dependem unicamente dos parâmetros $m_1 = (|(t-1) - (g-1)| - 1)2^{-1}$ e $m_2 = (n-g - (t-1)-1) 2^{-1}$. Os testes estatísticos associados à matriz de covariância (Σ) do modelo multivariado não consideram a condição H-F. Além disso, para obter o número de parâmetros da matriz (Σ) basta utilizar a expressão: $t(t+1)/2$, em que "t" corresponde ao número de medidas repetidas no tempo. As hipóteses testadas na análise multivariada foram as mesmas da análise univariada de perfil.

3.3. Análise via modelos mistos

A utilização dos modelos mistos teve como objetivo testar diversas estruturas de covariâncias para G e R, assim como diversos conjuntos de efeitos aleatórios. Para isso dividiu-se a análise via modelos mistos em 4 grupos de modelos (U, E, N e F) visando à seleção do melhor dentro de cada grupo e posteriormente a comparação entre os melhores de cada grupo. De maneira geral o modelo adotado foi:

$$Y = X\beta + Zb + \varepsilon$$

em que:

Y = é o vetor de observações ou vetor ($n \times 1$) de respostas da i-ésima unidade experimental ou indivíduo;

X = é uma matriz ($n \times p$) conhecida e de especificação, de posto completo dos efeitos fixos;

β = é um vetor ($p \times 1$) de parâmetros de efeitos fixos (tempo) desconhecidos;

$Z =$ é uma matriz ($n \times q$) conhecida e de especificação, de posto completo dos efeitos aleatórios;

$b =$ é um vetor ($q \times 1$) de parâmetros de efeitos aleatórios (intercepto, tempo e tempo²) desconhecidos, distribuído como $N(0, G)$, sendo a matriz G positiva definida de dimensão $q \times q$;

$\varepsilon =$ é o vetor ($n \times 1$) de erros aleatórios não observáveis, associado à unidade experimental i , distribuído como $N(0, R)$, ou seja, distribuído normalmente com média 0 e matriz de covariância R , sendo a matriz R positiva definida de dimensão $t \times t$.

Em todos os modelos gerados foram testadas as seguintes estruturas de covariâncias: não estruturada (UN), simetria composta heterogênea (CSH), autorregressiva de primeira ordem heterogênea (ARH(1)), toeplitz (TOEP), simetria composta (CS), componente de variâncias (VC), diagonal principal banded (UN(1)), toeplitz banded (TOEP2) e autorregressiva de primeira ordem (AR(1)).

Nesta primeira parte da análise foram considerados 4 grupos de modelos mistos (U, E, N e F), sendo todos com a mesma estrutura fixa. O conjunto de efeitos fixos foi formado pelo tempo, tempo², genótipos e interações. As estruturas de covariâncias para G (entre indivíduos, efeitos aleatórios) e R (intra-indivíduos) utilizadas foram as descritas acima. Os grupos formados foram:

Grupo U: neste grupo o conjunto de efeitos aleatórios foi formado pelo Intercepto, tempo e tempo²;

Grupo E: para este grupo retirou-se o efeito aleatório de intercepto;

Grupo N: no grupo N considerou-se como efeito aleatório somente o efeito de tempo²;

Grupo F: este grupo possui como efeito aleatório somente o erro, ou seja, a variação intra - indivíduo (R).

Inicialmente, em cada grupo, considerou-se a parte fixa (efeitos fixos) e a parte aleatória definida para cada grupo e variaram-se as estruturas de

4.0 RESULTADOS E DISCUSSÃO

4.1 Análise Exploratória

Nas figuras 1 e 2 são apresentados os perfis individuais e médios dos genótipos de feijão vagem avaliados para a variável altura de plantas. Segundo Ker (2010), o gráfico de perfis individuais e médios é uma das representações gráficas mais importantes na análise descritiva longitudinal. Dentre os objetivos deste tipo de representação gráfica pode destacar: identificar observações discrepantes, comparar as variabilidades de cada grupo, sugerir possíveis modelos para explicar a variação da trajetória média ao longo do tempo.

Analisando-se as figuras 1 e 2 não são verificados perfis nem observações discrepantes. Observa-se que os perfis individuais dos genótipos são bastante homogêneos, com exceção dos genótipos Top Seed Blue Line, UENF 7-4-1 e UENF 14-4-3 (Figura 1). Além disso, supõe-se que um polinômio do segundo grau possa explicar bem o crescimento dos genótipos de feijão vagem (Figura 2).

A altura de plantas apresentou uma resposta crescente durante todo o período avaliado, indicando efeito significativo do tempo, ou seja, a possível existência de perfis não horizontais (Figuras 1 e 2). No entanto, nota-se que o crescimento inicial é pequeno e bastante semelhante entre os genótipos e com o decorrer do tempo observa-se maiores diferenças entre os genótipos, evidenciando o efeito de genótipo, ou seja, de que os perfis possam ser não coincidentes. Além disso, a interação genótipo x tempo tende a ser significativa devido à ausência de paralelismo entre os perfis médios (Figura 2).

4.2 Análise univariada de perfil

Os resultados da análise univariada de perfil da variável altura de planta considerando o esquema de parcela subdividida no tempo indicam efeito significativo de genótipos, tempo e da interação genótipo x tempo (Tabela 2). Observa-se que a hipótese de perfis coincidentes, testada para o fator entre genótipos foi rejeitada ($p < 0,0001$) pelo teste F, indicando que a altura do feijão vagem apresenta comportamento diferenciado conforme o genótipo considerado.

Tabela 2. Resultado da análise univariada de perfil para os dados da variável altura de planta de 17 genótipos de feijão vagem (*Phaseolus vulgaris*. L) avaliados em 9 medidas no tempo.

Causas da variação	Graus de liberdade	Soma de quadrados	Quadrado médio	Teste F	(Pr>F)
Modelo	203	2240609,385	11037,485	222,14	<, 0001
Genótipo (G)	16	144869,627	9054,352	27,47	<, 0001
Bloco	3	5083,821	1694,607		
Erro a	48	15822,129	329,628	6,63	
Tempo (T)	8	1933830,18	241728,773	4864,97	<, 0001
G x T	128	141003,621	1101,591	22,17	<, 0001
Erro b	408	20272,562	49,688		
Total	611	2260881,947			

R²=0,9910; Média da Altura de planta (cm) = 57,87
CV parcela=31,37% e CV subparcela=12,18%

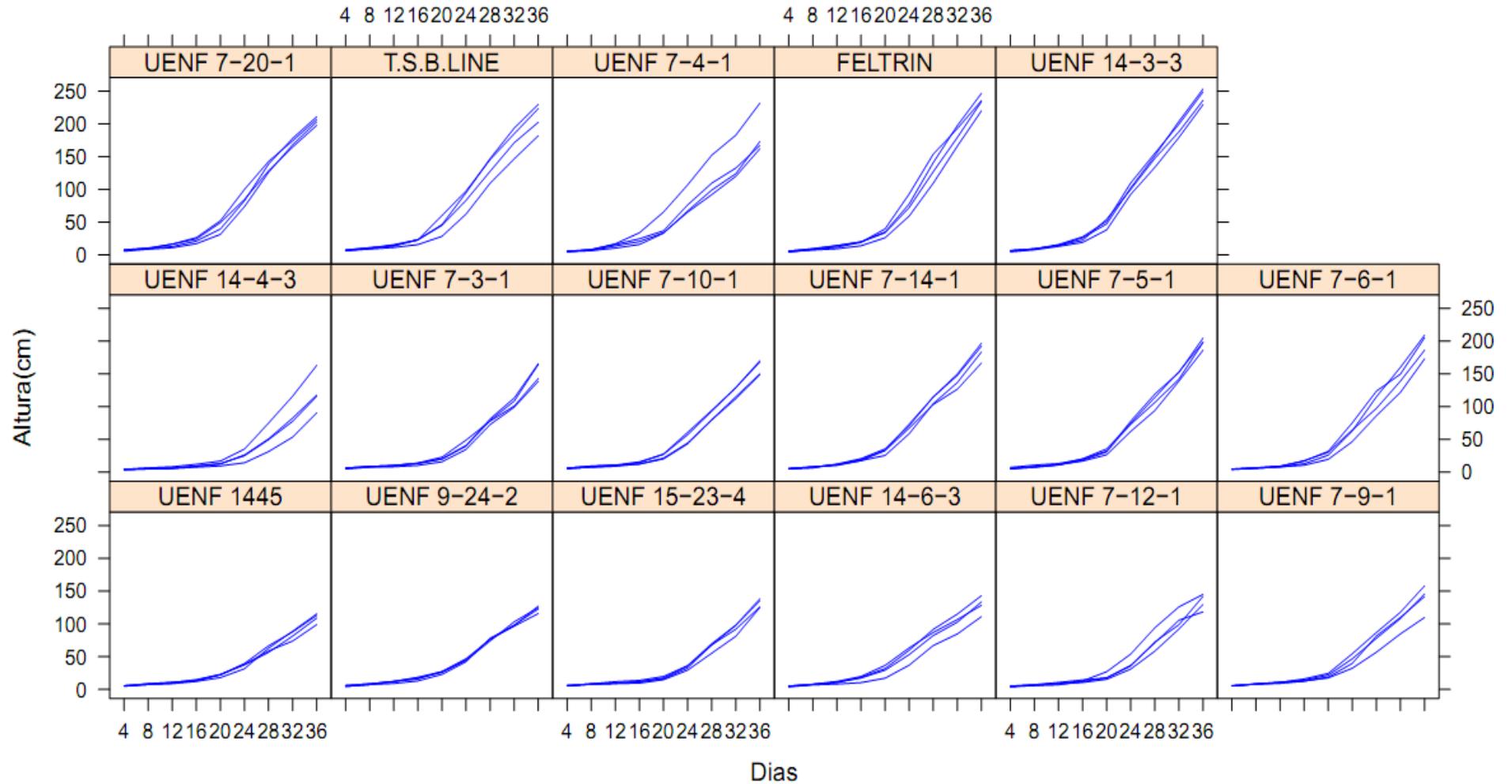


Figura 1. Representação gráfica dos perfis individuais para a variável altura de plantas de 17 genótipos de feijão vagem, em função dos tempos de avaliação.

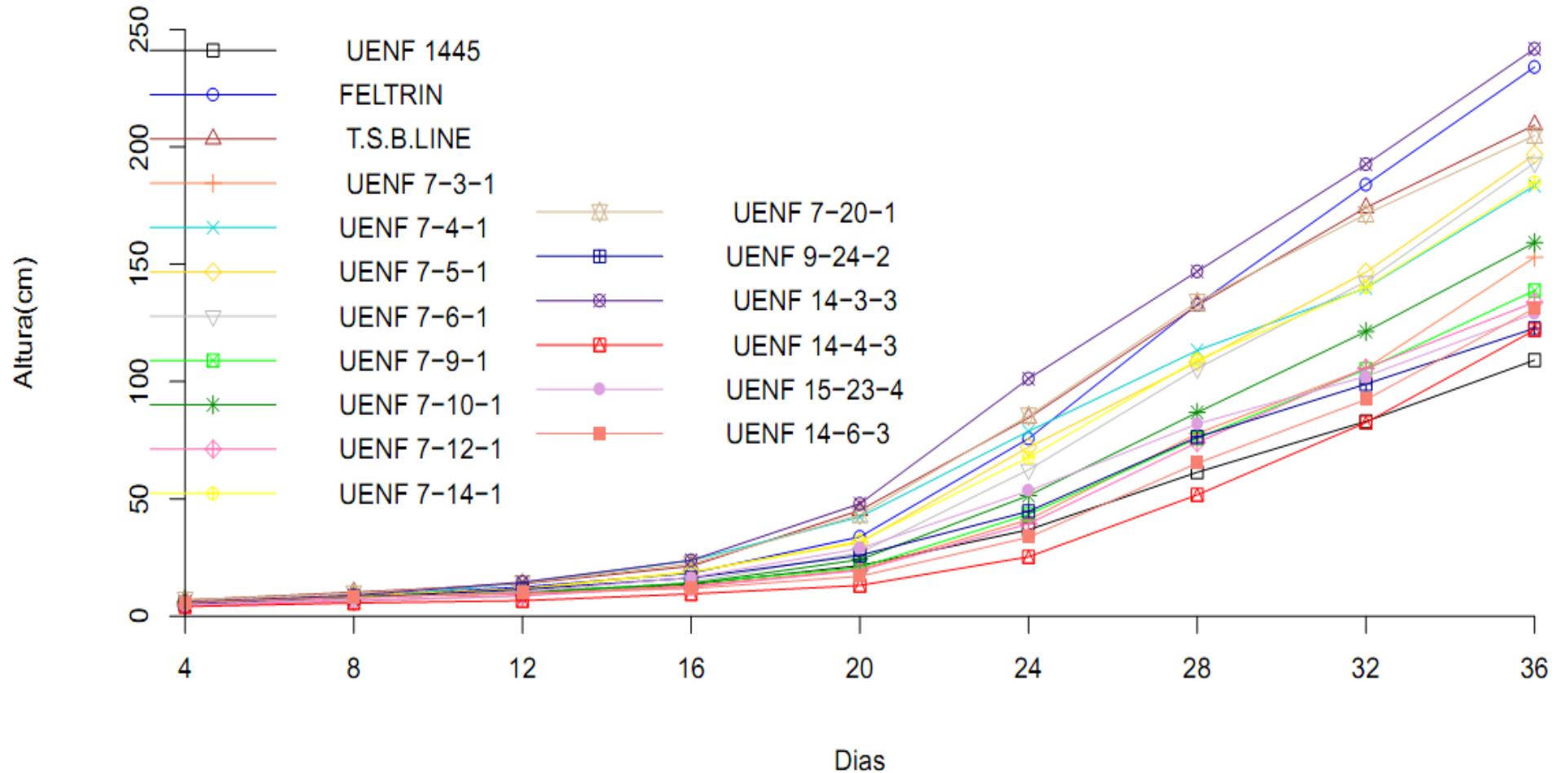


Figura 2. Representação gráfica dos perfis médios para a variável altura de plantas de 17 genótipos de feijão vagem ao longo o tempo.

Como esperado houve rejeição pelo teste F ($p < .0001$) da hipótese de perfis horizontais, que testa o efeito do tempo, indicando que à medida que o tempo progride há variação no crescimento em altura do feijão vagem. De forma semelhante, a hipótese de perfis paralelos também foi rejeitada ($p < .0001$), indicando que a interação genótipo x tempo atua de forma significativa no crescimento em altura do feijão vagem. Os resultados obtidos ratificam os observados na análise exploratória.

Em estudos de medidas repetidas no tempo, em um delineamento no esquema de parcelas subdivididas, por exemplo, os níveis do fator tempo não podem ser aleatorizados. Dessa forma, a análise de variância usual pode não ser válida, porque com a falta de aleatorização, os erros correspondentes a respectivas unidades experimentais ou indivíduos podem ter uma matriz de covariâncias que não é igual àquela exigida para que a análise usual de um delineamento seja válida, isto é, variâncias homogêneas (Xavier & Santos, 2001). Além disso, a falta de aleatorização pode inflacionar a probabilidade de falsamente rejeitar a hipótese nula (erro tipo 1) (Fernandez, 1991).

Para que a análise de dados longitudinais considerando o esquema de parcelas subdivididas no tempo, em nível de subparcela, para o fator tempos e interação tempos x tratamentos, seja válida, a matriz de covariâncias deve apresentar a forma chamada HUYNH-FELDT (H-F) (variâncias iguais e correlações nulas), o que pode ser verificado através do teste de esfericidade ou circularidade proposto por Mauchly (1940). O resultado do teste de Mauchly para o presente trabalho é apresentado na tabela 3.

Tabela 3. Resultado do teste de esfericidade de Mauchly

Graus de liberdade de F	Critério de Mauchly (W)	Estatística Qui – quadrado X^2	Valor - p
35	1,5724E-8	810,811	$<<, 0001$

De acordo com o teste de Mauchly, realizado com 35 graus de liberdade e aproximação de *Qui-Quadrado* (810,811) a condição de esfericidade foi rejeitada ($p < 0,001$). Isto significa que a matriz de variância-covariância não pode ser considerada do tipo H-F, uma vez que a suposição de esfericidade não foi atendida e, portanto o teste univariado não deve ser utilizado.

Avaliando o crescimento em altura ao longo do tempo da *Leucaena leucocephala* (Lam.) de Wit Souza et al. (2008) também verificaram rejeição da condição de esfericidade pelo teste de esfericidade de Mauchly.

Aparício et al. (2010) avaliaram em experimento, o crescimento inicial de dois clones do eucalipto e verificaram que o teste rejeitou a hipótese nula de esfericidade da matriz de covariância e, por esse fato, optaram pela não realização da análise univariada de perfil.

Ferreira (2012) avaliando o efeito de diferentes dietas com e sem ingestão de café no ganho de peso de animais (ratos) ao longo do tempo também observou que a condição de esfericidade foi rejeitada pelo teste de Mauchly. Resultados semelhantes foram obtidos por Freitas et al. (2007). Por outro lado, trabalhando com dados longitudinais em experimentação animal Cestari et al. (2012) observaram que a condição de esfericidade não foi violada pelo teste de Mauchly. Neste caso, o teste univariado poderia ser utilizado, pois a matriz de variância-covariância apresenta variâncias comuns e covariâncias nulas.

Uma alternativa para a rejeição da suposição de esfericidade é a correção do número de graus de liberdade para o fator tempo e a interação genótipo x tempo, ou seja, mesmo que a condição de esfericidade não seja satisfeita a distribuição F central pode ser utilizada de forma aproximada (Box, 1954 a e b; Geisser & Grenhouse, 1958; Huynh & Feldt, 1976).

Os resultados da análise univariada de perfil da variável altura de plantas com a correção dos números de graus de liberdade considerando o esquema de parcela subdividida no tempo para os fatores da subparcela (intra-indivíduos) são apresentados na tabela 4.

Tabela 4. Resultado da análise univariada de perfil com correção dos números de graus de liberdade para os dados da variável altura de planta de 17 genótipos de feijão vagem (*Phaseolus vulgaris*. L) avaliados em 9 medidas no tempo.

CV	GL	SQ	QM	F	(Pr>F)	G-G	H-F
T	8	1933828,614	241728,577	4865,02	<, 0001	<, 0001	<, 0001
G x T	128	141003,892	1101,593	22,17	<, 0001	<, 0001	<, 0001
Erro b	408	20272,309	49,687				

G-G= Greenhouse-Geisser $\epsilon = 0,1857$

H-F= Huynh-Feldt $\epsilon = 0,2501$

CV= Causa da variação; SQ=Soma de quadrado; QM=Quadrado médio; F= Teste F; G=Genótipos; T=Tempo.

As correções dos graus de liberdade não promoveram mudanças nos resultados obtidos para os efeitos de tempo e da interação genótipo x tempo (Tabela 4).

Trabalhando com dados longitudinais na cultura da cana-de-açúcar Freitas et al. (2008) também observaram que as correções G-G e H-F não alteraram as significâncias do teste F para os efeitos de tempo e interação genótipo x tempo. Resultados semelhantes foram obtidos por Xavier (2000), Freitas (2007), Amador (2010), Ferreira (2012) e Santos et al. (2013). Levando-se em consideração os resultados obtidos no presente trabalho e os observados na literatura, as correções dos graus de liberdade não parecem ser muito efetivas na maioria das situações, apesar de ser uma alternativa na análise de dados longitudinais.

Apesar de no presente trabalho não ter sido possível detectar diferenças entre as correções utilizadas, Freitas et al. (2011) trabalhando cortes sequenciais na cultura da alfafa verificaram que as correções não alteraram os valores da probabilidade F para o efeito de corte, mas a interação 'tratamento x corte' foi não significativa para a correção G-G e significativa para H-F. Resultados obtidos por Malheiros (2004) indicam como melhor alternativa a correção G-G, assim como também sugerem Littell et al. (1998).

Malheiros (2004) por meio de simulações concluiu que tanto para tempos como para a interação tempo x tratamento, com a correção de Huynh & Feldt (1976) os testes F da análise da variância são imprecisos e com a correção de Greenhouse & Geisser (1959) são precisos, independente dos dados serem balanceados ou não e da estrutura da matriz Σ . Quando não se faz qualquer correção, os testes F para a interação tempo x tratamento apresentaram-se precisos e, para tempos, a precisão depende da estrutura da matriz Σ .

Em virtude da matriz de covariâncias não poder ser considerada do tipo HF e as correções não terem promovido nenhuma mudança nos resultados indicados pela análise univariada de perfil considerou-se o modelo multivariado, que não impõe nenhuma restrição quanto a matriz de covariâncias do erro.

4.3 Análise multivariada de perfil

A matriz abaixo apresenta as variâncias (diagonal), covariâncias (diagonal superior) e correlações (diagonal inferior) considerando o modelo multivariado.

	T1	T2	T3	T4	T5	T6	T7	T8	T9
T1	15,809	14,405	17,341	19,425	43,558	28,253	12,580	-5,369	21,718
T2	0,7117	25,908	30,345	45,633	88,855	114,01	140,16	148,86	171,47
T3	0,4826	0,6596	81,685	123,12	240,71	285,30	311,10	303,76	311,03
T4	0,2706	0,4966	0,7546	325,90	669,36	940,03	1069,2	1078,9	1050,2
T5	0,2513	0,4005	0,6110	0,8506	1900,2	2452,6	2749,3	2601,9	2447,6
T6	0,1107	0,3490	0,4919	0,8114	0,8768	4018,1	4933,2	4679,3	4462,8
T7	0,0371	0,3231	0,4039	0,6950	0,7401	0,9020	7263,3	7276,9	7311,6
T8	-0,0147	0,3178	0,3652	0,6994	0,6486	0,7924	0,9279	8468,0	8696,5
T9	0,0535	0,3299	0,3370	0,5696	0,5498	0,6809	0,8400	0,9253	10430,7

Como pode ser observado na matriz acima, as variâncias são heterogêneas e crescentes ao longo tempo e as covariâncias são não nulas com comportamento crescente até a quinta avaliação, diminuindo a partir desse ponto. Além disso, observa-se que os dados são correlacionados sendo, que medidas tomadas em tempos mais próximos são mais fortemente correlacionadas. Este resultado concorda com os obtidos por Littell et al. (1998), Xavier e Santos (2001), Costa (2003), Rosário (2003) e Ferreira (2012), que também verificaram correlações maiores em tempos mais próximos, porém discorda dos observados por Freitas (2007), Barbosa (2009), Amador (2010) e Silva (2011).

Por meio dos testes multivariados testaram-se as mesmas hipóteses estudadas na análise univariada de perfil, ou seja, de perfis horizontais, paralelos e coincidentes. Pode-se observar na tabela 5 que a hipótese de perfis horizontais foi rejeitada por todos os testes multivariados utilizados (Wilks'Lambda, Pillai's Trace, Hotelling – Lawley Trace e Roy's Greatest Root), assim como as hipóteses de perfis paralelos e coincidentes. Assim como foi observado por Xavier (2000), Amador (2010) e Ferreira (2012) os resultados obtidos pela análise multivariada não se distinguiram dos observados na análise univariada de perfil.

Tabela 5. Resultado da análise multivariada de perfil para os dados da variável altura de planta de 17 genótipos de feijão vagem (*Phaseolus vulgaris*. L) avaliados em 9 medidas no tempo.

H= Matriz de soma de quadrados e produtos cruzados para o fator Tempo E= Matriz de soma de quadrados e produtos cruzados para os resíduos S=1 M=3 N=19,5					
Estatística	Valor	F	GL ¹	GL ²	Pr>F
Wilks' Lambda	0,00454983	1121,29	8	41	<,0001
Pillai's Trace	0,99545017	1121,29	8	41	<,0001
Hotelling – Lawley Trace	218,78835051	1121,29	8	41	<,0001
Roy's Greatest Root	218,78835051	1121,29	8	41	<,0001
H= Matriz de soma de quadrados e produtos cruzados para a interação Genótipo x Tempo E= Matriz de soma de quadrados e produtos cruzados para os resíduos S=8 M=3,5 N=19,5					
Estatística	Valor	F	GL ¹	GL ²	Pr>F
Wilks' Lambda	0,00058853	4,35	128	308,37	<,0001
Pillai's Trace	3,88237546	2,83	128	384	<,0001
Hotelling – Lawley Trace	26,13578574	8,05	128	178,12	<,0001
Roy's Greatest Root	18,26399722	54,79	16	48	<,0001

Em relação ao modelo univariado (análise univariada de perfil), a análise multivariada de perfil tem como principal vantagem o fato de permitir a adoção de um modelo bastante geral para representar a estrutura de variância-covariância, admitindo que as variâncias das respostas em cada tempo e as variâncias-covariâncias das respostas entre tempos distintos sejam diferentes (LIMA, 1996). Isto ocorre porque o modelo multivariado utiliza uma matriz de variâncias-variância-covariância sem restrições, chamada não estruturada (Scheiner & Gurevitch, 2001). No entanto, apresenta como desvantagem a necessidade de perfis completos e o baixo poder dos testes de hipóteses, devido à excessiva parametrização. Além disso, em situações em que o número de unidades experimentais é pequeno em relação ao número de ocasiões de avaliação a eficiência dos estimadores dos parâmetros pode ser afetada (Timm, 1980; Wald, 2000).

Outra desvantagem é que o modelo multivariado completo e não estruturado, tende a apresentar problemas de convergência, em virtude das altas correlações, geralmente verificadas entre medidas repetidas, e do elevado número de parâmetros a ser estimado. A abordagem multivariada torna-se proibitiva quando o número de colheitas ou safras é elevado. Algumas vezes,

mesmo com pequeno número de safras, as matrizes de covariância não estruturadas não são positivas definidas e conduzem a correlações maiores que 1 (Mariguele et al., 2011).

Medidas repetidas, altamente correlacionadas, aumentam o risco de obtenção de matrizes de covariância não positivas definidas e de não convergência, no processo de estimação de componentes de variância. Também, quanto maior a ordem da matriz de covariância, maior é a chance de que ela não seja positiva definida (Resende, 2007).

Perri et al. (1999) afirmam que uma abordagem mais atual consiste no uso de modelos lineares mistos, que se baseia em três aspectos fundamentais: estimação e teste de hipóteses sobre os parâmetros de efeito fixo, predição dos parâmetros de efeito aleatório e estimação dos componentes de variância.

Segundo Camarinha Filho (2002), o sucesso do procedimento de modelagem está fortemente associado ao exame dos efeitos aleatórios e à possibilidade de se introduzir, no modelo, estruturas de variâncias e covariâncias.

4.4. Modelo Misto

Os modelos mistos englobam a análise de curvas de crescimento ou curvas polinomiais, que levam em conta a matriz de covariâncias que melhor explica o comportamento das observações (Von Ende, 1993; Littell et al., 1998). Nesse sentido, foram gerados quatro grupos de modelos mistos, sendo cada um com um conjunto de efeitos aleatórios visando obter o modelo que melhor se ajusta aos dados.

Os modelos foram selecionados utilizando os critérios de AIC, BIC e $-2\text{Res log likelihood}$, sendo o critério mais importante o AIC e quando possível por meio do teste de razão de verossimilhanças.

4.4.1. Modelo Misto com efeitos aleatórios de intercepto, tempo e tempo²(U)

Na tabela 6 são apresentados as estruturas testadas e os valores dos critérios de informação obtidos para a seleção da melhor estrutura para G

(associada aos parâmetros de efeitos aleatórios) considerando o modelo maximal, com todos os parâmetros de efeitos fixos (tempo, tempo², genótipos e interações) e de efeitos aleatórios (intercepto, tempo e tempo²).

Tabela 6. Estruturas de variâncias-covariâncias (Σ) testadas para a matriz G (variação entre indivíduos), critérios de informação e número de parâmetros de covariâncias obtidos para a altura de genótipos de feijão vagem avaliada ao longo do tempo, considerando os efeitos de intercepto, tempo e tempo² como aleatórios.

Estruturas		Critérios de informação				NP ⁵
R	G	-2Res log Likelihood ¹	AIC ²	AICC ³	BIC ⁴	
VC	TOEP (2)	4208,0	4214,0	4214,1	4220,7	3
	AR (1)	4208,0	4214,0	4214,1	4220,7	3
	CS	4208,1	4214,1	4214,1	4220,7	3
	VC					
	CSH					
	UN					
	ARH (1)		Não convergiram			
	UN (1)		Não convergiram			

¹ Valor que maximiza o logaritmo da função de verossimilhança; ² Critério de informação de Akaike; ³ Critério de informação de Akaike corrigido; ⁴ Critério de informação Bayesiano; ⁵ Número de parâmetros de covariância do modelo.

Pode-se observar uma grande semelhança entre as estruturas que convergiram (Tabela 6). Não há diferença entre as estruturas toeplitz banded (TOEP(2)) e a estrutura autorregressiva de primeira ordem (AR(1)). Nesse caso, escolheu-se a estrutura AR(1) por ser uma estrutura em que as covariâncias diminuem com o tempo. Essa estrutura supõe homogeneidade de variâncias e heterogeneidade de covariâncias.

A Σ do tipo simetria composta (CS) foi a que apresentou a pior qualidade de ajuste dentre as estruturas estudadas, porém a diferença para as demais foi muito pequena. Essa estrutura supõe homogeneidade de variâncias dos efeitos aleatórios ao longo do tempo e homogeneidade de covariâncias para medidas realizadas em tempos diferentes.

As estruturas TOEP, CSH, UN, UN(1) e ARH não convergiram. Tais estruturas podem ser consideradas mais complexas quando comparadas com as que convergiram. Além disso, não foi possível obter para G uma matriz positiva definida, condição exigida para obtenção de estimativas confiáveis pelos métodos adotados nesse estudo.

Amador (2010) testando diferentes estruturas de covariâncias para o ajuste de curvas de crescimento de *Ricinus communis* L. ao longo do tempo, observou que não houve convergência para as Σ UN, autorregressiva de primeira ordem (AR (1)), simetria composta heterogênea (CSH), autorregressiva de primeira ordem heterogênea (ARH (1)) e diagonal principal banded (UN (1)). Somente houve convergência para a Σ componente de variância (VC). Já Ferreira et al. (2013) estudando a influência do uso de café no controle de pesos de animais (ratos) por meio do modelo misto e testando diferentes estruturas de covariâncias para a seleção de G obtiveram como melhor qualidade de ajuste a utilização da matriz UN para G.

A estimativa da Σ AR(1) para a matriz G (efeitos aleatórios) dos dados longitudinais em estudo é dada por:

$$\hat{\Sigma}AR(1) = \begin{pmatrix} 0,00017 & 0,00007 & 0,00003 \\ & 0,00017 & 0,00007 \\ sim & & 0,00017 \end{pmatrix}$$

Para a escolha da melhor estrutura para R (associada ao erro aleatório do modelo), ou seja, que represente a variação intra-indivíduos, fixou-se a estrutura selecionada para G (AR(1)) e testaram-se todas as demais estruturas. O ajuste dos dados foi possível somente quando se utilizou a estrutura do tipo componente de variâncias (VC) para R, que supõe que os erros sejam homocedásticos e independentes (covariâncias nulas). Resultado semelhante foi obtido por Amador (2010) e Ferreira et al. (2013).

Barbosa (2009) trabalhando com diferentes Σ em dados longitudinais de leite obtido pelo processo UAT (temperatura ultra-alta) observou que dentre as estruturas UN, CS e VC, esta última foi a que apresentou a melhor qualidade de ajuste para explicar a variabilidade intra-indivíduos.

Apesar dos resultados obtidos no presente trabalho e dos observados na literatura (Barbosa, (2009); Amador, (2010) e Ferreira et al. (2013)) apontaram para a modelagem de R com a estrutura VC, que teoricamente esta não é a estrutura mais adequada para a modelagem dos erros em dados longitudinais, visto que, nesse tipo de dado os erros tendem a ser heterocedásticos e correlacionados, característica que não é imposta pela estrutura VC. Por outro lado, quando se tem uma estrutura complexa para G pode se ter problemas de

modelagem ao se considerar estruturas mais complexas para R porque nesse caso há um aumento no número de parâmetros a serem estimados, o que dificulta a convergência na estimação dos componentes de variância.

A estimativa da Σ VC para a matriz R dos dados longitudinais é dada abaixo.

$$\hat{\Sigma}VC = \begin{pmatrix} 38,74 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & 38,74 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & 38,74 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & 38,74 & 0 & 0 & 0 & 0 & 0 \\ & & & & 38,74 & 0 & 0 & 0 & 0 \\ & & & & & 38,74 & 0 & 0 & 0 \\ & & & & & & 38,74 & 0 & 0 \\ & & & & & & & 38,74 & 0 \\ & & & & & & & & 38,74 \end{pmatrix}$$

sim

Para as demais estruturas ou não houve convergência do processo iterativo ou o processo foi parado por causa de muitas avaliações de verossimilhança. Neste último caso, acredita-se que o motivo seja o número elevado de efeitos aleatórios. Resultado semelhante foi obtido por Ferreira et al. (2013).

O modelo mais adequado para os dados em estudo considerando como efeitos aleatórios o intercepto, tempo e tempo² é aquele com Σ AR(1) para G e VC para R.

4.4.2. Modelo Misto com efeitos aleatórios de tempo e tempo²(E)

Nesse segundo grupo (E) de modelos gerados foi desconsiderado o efeito aleatório de intercepto e os resultados obtidos podem ser observados na tabela 7.

Tabela 7. Estruturas de variâncias-covariâncias (Σ) testadas para a matriz G (variação entre indivíduos), critérios de informação e número de parâmetros de covariâncias obtidos para a altura de genótipos de feijão vagem avaliada ao longo do tempo, considerando os efeitos de tempo e tempo² como aleatórios.

Estruturas		Critérios de informação				NP ⁵
R	G	-2Res log Likelihood ¹	AIC ²	AICC ³	BIC ⁴	
	AR (1)	4208,0	4214,0	4214,1	4220,7	3
	TOEP	4208,0	4214,0	4214,1	4220,7	3
	TOEP (2)	4208,0	4214,0	4214,1	4220,7	3
	CS	4208,0	4214,0	4214,1	4220,7	3
VC	UN (1)					
	VC					
	UN		Não convergiram			
	CSH		Não convergiram			
	ARH (1)		Não convergiram			

¹ Valor que maximiza o logaritmo da função de verossimilhança; ² Critério de informação de Akaike; ³ Critério de informação de Akaike corrigido; ⁴ Critério de informação Bayesiano; ⁵ Número de parâmetros de covariância do modelo.

Para esse grupo de modelos pode-se observar que os valores dos critérios de informação foram idênticos para as estruturas AR(1), TOEP(2), TOEP e CS (Tabela 7). Estas estruturas apresentam em comum o fato de possuírem variâncias homogêneas. Apesar de não existir diferenças entre tais estruturas quanto aos critérios de informação considerados, escolheu-se a estrutura AR(1) para representar a variação entre indivíduos por ser uma estrutura em que as covariâncias diminuem com o tempo. As estruturas CS, UN(1), VC, UN, CSH e ARH(1) não convergiram.

Trabalhando com o mesmo conjunto de efeitos aleatórios (tempo e tempo²) Amador (2010) observou que os melhores resultados para G foram obtidos com as estruturas VC e UN (1).

A estimativa da Σ AR(1) para o conjunto de efeitos aleatórios formados pelo tempo e tempo² (G) é dada por:

$$\hat{\Sigma}_{AR(1)} = \begin{pmatrix} 0,00017 & 0,00017 \\ sim & 0,00012 \end{pmatrix}$$

Considerando a estrutura autorregressiva de primeira ordem (AR(1)) para G e variando-se as demais estruturas para R, observou-se que não houve

convergência do método adotado para as estruturas testadas, com exceção da estrutura VC. A estimativa da Σ (VC) para o erro aleatório (R) do modelo é dada por:

$$\hat{\Sigma}_{VC} = \begin{pmatrix} 38,74 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & 38,74 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & 38,74 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & 38,74 & 0 & 0 & 0 & 0 & 0 \\ & & & & 38,74 & 0 & 0 & 0 & 0 \\ & & & & & 38,74 & 0 & 0 & 0 \\ & & & & & & 38,74 & 0 & 0 \\ & & & & & & & 38,74 & 0 \\ sim & & & & & & & & 38,74 \end{pmatrix}$$

4.4.3. Modelo Misto com efeito aleatório de tempo²(N)

Nesse terceiro grupo de modelos gerados considerou-se somente o efeito de tempo² como aleatório e os resultados podem ser visualizados na tabela 8.

Tabela 8. Estruturas de variâncias-covariâncias (Σ) testadas para a matriz G (variação entre indivíduos) e critérios de informação obtidos para a variável altura de plantas de 17 genótipos de feijão vagem avaliados ao longo do tempo, considerando o efeito de tempo² como aleatório.

Estruturas		Critérios de informação				NP ⁵
R	G	-2Res log Likelihood ¹	AIC ²	AICC ³	BIC ⁴	
	VC	4208,1	4212,1	4212,1	4216,5	2
	UN	4208,1	4212,1	4212,1	4216,5	2
	UN (1)	4208,1	4212,1	4212,1	4216,5	2
	TOEP	4208,1	4212,1	4212,1	4216,5	2
VC	TOEP (2)	4208,1	4212,1	4212,1	4216,5	2
	AR (1)					
	CSH					
	ARH (1)					
	CS					
			Não convergiram			

¹ Valor que maximiza o logaritmo da função de verossimilhança; ² Critério de informação de Akaike; ³ Critério de informação de Akaike corrigido; ⁴ Critério de informação Bayesiano; ⁵ Número de parâmetros de covariância do modelo.

Quando se considerou somente o efeito de tempo² como aleatório pôde-se observar que as estruturas foram semelhantes nos critérios de informação obtidos

e no número de parâmetros de covariâncias gerados (Tabela 8). Não houve conversão para as estruturas TOEP (2), AR(1), CSH, ARH(1) e CS. Dentre as estruturas que convergiram escolheu-se a estrutura com suposição mais simples. Nesse caso, a estrutura VC foi escolhida para representar a variabilidade entre indivíduos, ou seja, a matriz G. Sua estimativa é mostrada abaixo:

$$\hat{\Sigma}_{VC} = (0,0002)$$

Amador (2010) testando diferentes estruturas de covariâncias para o ajuste de curvas de crescimento de *Ricinus communis* L, ao longo do tempo e considerando o tempo² como efeito aleatório, observou que a melhor qualidade de ajuste foi obtida quando se utilizou a estrutura UN (1).

No próximo passo da análise fixou-se a estrutura VC para G e variaram-se as estruturas para R. Os resultados são apresentados na tabela 9.

Tabela 9. Estruturas de variâncias-covariâncias (Σ) testadas para a matriz R (variação intra-indivíduos) e critérios de informação obtidos para a variável altura de plantas de 17 genótipos de feijão vagem avaliados ao longo do tempo, considerando o efeito de tempo² como aleatório.

Estruturas		Critérios de informação				NP ⁵
R	G	-2Res log Likelihood ¹	AIC ²	AICC ³	BIC ⁴	
UN (1)		3904,0	3924,0	3924,4	3946,2	46
TOEP (2)		4002,1	4008,1	4008,2	4014,8	3
VC		4208,1	4212,1	4212,1	4216,5	2
AR (1)		4113,3	4119,3	4119,3	4125,9	3
TOEP	VC					
UN						
CSH			Não convergiram			
ARH (1)						
CS						

¹ Valor que maximiza o logaritmo da função de verossimilhança; ² Critério de informação de Akaike; ³ Critério de informação de Akaike corrigido; ⁴ Critério de informação Bayesiano; ⁵ Número de parâmetros de covariância do modelo.

Como pode ser observado, a estrutura que melhor se ajustou aos dados foi a estrutura UN (1) seguida das estruturas TOEP (2), AR (1) e VC. Mais uma vez para R a estrutura mais complexa foi a estrutura que melhor explicou a

variabilidade intra-indivíduos. As demais estruturas não convergiram. Vale ressaltar que os melhores resultados foram observados para estruturas (UN (1) e TOEP (2)) que apresentam variâncias heterogêneas.

A estimativa da matriz do tipo diagonal principal banded (UN (1)) para o erro aleatório (R) do modelo é dada por:

$$\hat{\Sigma}_{UN(1)} = \begin{pmatrix} 0,460 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & 9,417 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & 3,077 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & 41,87 & 0 & 0 & 0 & 0 & 0 \\ & & & & 147,5 & 0 & 0 & 0 & 0 \\ & & & & & 84,60 & 0 & 0 & 0 \\ & & & & & & 51,04 & 0 & 0 \\ & & & & & & & 12,32 & 0 \\ sim & & & & & & & & 119,4 \end{pmatrix}$$

Dentro das estruturas que convergiram, o pior resultado foi observado para estrutura VC.

Considerando os resultados obtidos pelo teste de razão de verossimilhanças, observa-se que a superioridade apontada pelos critérios de informação da estrutura UN (1) em relação às demais foi confirmada (Tabela 10). O mesmo foi observado para as outras comparações realizadas, ou seja, as estruturas TOEP (2) e AR (1) demonstraram ser superiores à estrutura VC.

Tabela 10. Comparação de modelos com diferentes estruturas de variâncias-covariâncias (Σ), tomados dois a dois, segundo o teste de razão de verossimilhança (χ^2) para a variável altura de plantas de 17 genótipos de feijão vagem avaliados ao longo do tempo, considerando o efeito de tempo² como aleatórios.

Comparação	χ^2	GL	P-VALOR
UN (1) x TOEP (2)	98,10	43	<0,00001
UN (1) x VC	304,1	44	<0,00001
UN (1) x AR (1)	209,3	43	<0,00001
TOEP (2) x VC	206,0	1	<0,00001
VC x AR (1)	94,80	1	<0,00001

¹ Valor que maximiza o logaritmo da função de verossimilhança; ² Critério de informação de Akaike; ³ Critério de informação de Akaike corrigido; ⁴ Critério de informação Bayesiano; ⁵ Número de parâmetros de covariância do modelo.

4.4.4. Modelo Misto sem a variação entre indivíduos (G) (F).

Neste último grupo de modelos gerados desconsiderou-se os efeitos aleatórios correspondentes às inclinações das curvas para cada indivíduo. Logo, o objetivo foi escolher a melhor estrutura para R. Os resultados são apresentados na tabela 11.

Tabela 11. Estruturas de variâncias-covariâncias (Σ) testadas para a matriz R (variação intra-indivíduos) e critérios de informação obtidos para a variável altura de plantas de 17 genótipos de feijão vagem avaliados ao longo do tempo, sem efeitos aleatórios.

Estruturas	Critérios de informação					
	R	-2Res log Likelihood ¹	AIC ²	AICC ³	BIC ⁴	NP ⁵
UN		3270,1	3360,1	3368,1	3460	45
ARH (1)		3760,6	3780,6	3781	3802,8	10
TOEP		4033,4	4051,4	4051,8	4071,4	9
AR (1)		4172,8	4176,8	4176,9	4181,3	2
TOEP (2)		4203,1	4207,1	4207,1	4211,5	2
CS		4469,2	4473,2	4473,2	4477,6	2
VC		4588,2	4590,2	4590,2	4592,4	1
UN (1)						
CSH			Não convergiram			

¹ Valor que maximiza o logaritmo da função de verossimilhança; ² Critério de informação de Akaike; ³ Critério de informação de Akaike corrigido; ⁴ Critério de informação Bayesiano; ⁵ Número de parâmetros de covariância do modelo.

Como pode ser observado na tabela 11 a estrutura que melhor se ajustou aos dados foi a estrutura UN, seguida das estruturas ARH (1), TOEP, AR (1), TOEP (2), CS e VC. Os melhores ajustes foram observados para estruturas que permitiam a imposição de variâncias e covariâncias heterogêneas (UN (1) e ARH (1)). Por outro lado, os piores ajustes (VC e CS) foram observados para estruturas que impõem variâncias e covariâncias homogêneas, sejam elas nulas (VC) ou não (CS).

Barbosa (2009) trabalhando com modelos mistos no ajuste do crescimento de frangos de corte da linhagem Hubbard observou que dentre as estruturas testadas a UN foi a que apresentou o menor valor para o critério de AIC. Contudo, o autor decidiu pela estrutura ARH (1) que possuía menor número de parâmetros e menor valor de BIC.

Cestari et al. (2012) trabalhando com dados longitudinais em experimentação animal e testando diferentes estruturas de covariância para a modelagem adequada para R observaram que dentre as estruturas não estruturada (UN), Huynh-Feldt (HF), Autorregressiva de ordem 1 (AR (1)) e Simetria Composta (CS) o melhor resultado foi obtido para a estrutura AR (1).

Analisando o resultado do teste de razão de verossimilhanças (Tabela 12) observa-se que os resultados obtidos pelos critérios de informação foram confirmados. Vale ressaltar, que a estrutura VC foi a que apresentou o pior resultado, não sendo indicada para representar a variabilidade entre-indivíduos. Resultado esperado, pois em experimentos em que as avaliações são realizadas nas mesmas unidades experimentais ao longo do tempo é de esperar que ocorra uma correlação entre as medidas no tempo e exista certa heterogeneidade de variâncias (Littell et al., 2006).

Tabela 12. Comparação de modelos com diferentes estruturas de variâncias-covariâncias (Σ), tomados dois a dois, segundo o teste de razão de verossimilhança (χ^2) para a variável altura de plantas de 17 genótipos de feijão vagem avaliados ao longo do tempo, sem efeito aleatório.

Comparação	χ^2	GL	P-VALOR
UN x ARH (1)	490,50	35	<0,00001
UN x TOEP	763,30	36	<0,00001
UN x AR (1)	902,70	43	<0,00001
UN x TOEP (2)	933,00	43	<0,00001
UN x CS	1199,1	43	<0,00001
UN x VC	1318,1	44	<0,00001
ARH (1) x TOEP	272,80	1	<0,00001
ARH (1) x AR (1)	412,20	8	<0,00001
ARH (1) x TOEP (2)	442,50	8	<0,00001
ARH (1) x CS	708,60	8	<0,00001
ARH (1) x VC	827,60	9	<0,00001
TOEP x AR (1)	139,40	7	<0,00001
TOEP x TOEP (2)	169,70	7	<0,00001
TOEP x CS	435,80	7	<0,00001
TOEP x VC	554,80	8	<0,00001
AR (1) x VC	415,40	1	<0,00001
TOEP (2) x VC	385,10	1	<0,00001
CS x VC	119	1	<0,00001

A estimativa da matriz do tipo não estruturada (UN) para o erro aleatório (R) do modelo é dada por:

$$\hat{\Sigma}_{UN} = \begin{pmatrix} 2,74 & 0,17 & 4,66 & 12,95 & 17,81 & 2,58 & -16,07 & -20,17 & -24,07 \\ & 0,62 & 0,62 & 0,62 & 1,56 & 3,85 & 5,28 & 5,56 & 6,33 \\ & & 10,84 & 27,83 & 42,51 & 22,18 & -11,08 & -19,65 & -30,85 \\ & & & 77,70 & 117,35 & 55,41 & -39,89 & -62,56 & -94,76 \\ & & & & 199,99 & 133,98 & 8,02 & -27,59 & -89,44 \\ & & & & & 221,71 & 223,84 & 203,70 & 151,94 \\ & & & & & & 418,62 & 442,06 & 428,15 \\ & & & & & & & 506,87 & 502,30 \\ & & & & & & & & 570,94 \end{pmatrix}$$

sim

4.4.5. Comparação dos modelos selecionados dentro de cada grupo (UxExNx F) formado

Os valores para os critérios de informação obtidos para os modelos selecionados dentro de cada grupo (UxExNx F) são apresentados a seguir:

Tabela 13. Relação e comparação dos modelos lineares mistos selecionados dentro de cada grupo (UxExNx F) para a variável altura de plantas de 17 genótipos de feijão vagem avaliados ao longo do tempo.

Modelos	Estruturas		Efeito Fixo	Efeito Aleatório	-2 log Veross ¹	AIC ²	BIC ³
	R	G					
U	VC	AR(1)		Intercepto, tempo e tempo ² ,	4208,0	4214,0	4214,1
E	VC	AR(1)	Tempo, tempo ² ,	Tempo e tempo ²	4208,0	4214,0	4220,7
N	UN (1)	VC	genótipos e interações,	Tempo ² Sem a variação entre - indivíduos (G)	3904,0	3924,0	3946,2
F	UN	-			3270,10	3360,1	3460,0

¹ Valor que maximiza o logaritmo da função de verossimilhança; ² Critério de informação de Akaike; ³ Critério de informação Bayesiano.

Como se pode observar na tabela 13 o melhor modelo para o conjunto de dados longitudinais em estudo representado pelos menores valores para os critérios de informação foi o formado por todos os efeitos fixos (tempo, tempo², genótipos e interações), porém sem a matriz G (variação entre indivíduos), sendo que a matriz covariâncias mais adequada para os erros aleatórios do modelo foi a

do tipo não estruturada (UN). Essa estrutura permite a existência de diferentes covariâncias.

A matriz G consiste na matriz de variâncias e covariâncias do vetor b (vetor de parâmetros de efeitos aleatórios do modelo). Este por sua vez representa a diferença entre a curva do tratamento e a curva de cada indivíduo e são inseridos no modelo com o objetivo de proporcionar um melhor ajuste (Xavier, 2000; Amador, 2010). No entanto, para o conjunto de dados em questão a inserção de efeitos aleatórios (matriz G) mesmo considerando diferentes estruturas de covariâncias não melhorou a qualidade de ajuste comparando-se ao modelo sem a matriz G, visto que quando se considerou os efeitos aleatórios (matriz G) obtiveram-se os maiores valores para os critérios de informação (Tabela 13).

Considerando os modelos selecionados com efeitos aleatórios (matriz G) pode-se observar que o intercepto não melhorou a qualidade de ajuste. Este resultado era esperado porque como pode ser observado na figura 2, na análise exploratória, o crescimento dos genótipos estudados praticamente se inicia no mesmo ponto, ou seja, há pouca variabilidade inicial entre os diversos genótipos estudados.

Comparando o modelo selecionado do grupo U com o do grupo E observa-se praticamente que não há diferenças entre eles. A diferença foi observada somente quando se considerou o critério de informação Bayesiano (BIC). Contudo, quando se considerou somente o efeito de tempo² como efeito aleatório melhorou-se a qualidade de ajuste, mas ainda assim foi bem inferior ao ajuste dos dados considerando a matriz de covariâncias não estruturada para R e sem efeitos aleatórios.

Os resultados dos testes estatísticos (p-valor) para os efeitos fixos dos melhores modelos selecionados dentro de cada grupo UxExNxF são apresentados na tabela 14. Pode-se observar que os diferentes modelos selecionados não influenciaram na conclusão sobre os efeitos fixos, ou seja, em todos os modelos os efeitos fixos foram significativos e, portanto devem permanecer no modelo.

Tabela 14. Resultados (p-valores) de testes de significância (F-Snedecor) para os efeitos fixos dos modelos selecionados dentro de cada grupo UxExNxF para a variável altura de plantas de 17 genótipos de feijão vagem ao longo do tempo.

Fonte de variação	Modelos Selecionados			
	U	E	N	F
G	0,0168	<, 0046	<, 0001	<, 0001
T	<,0001	<, 0001	<,0001	<,0001
T*G	0,0001	<,0001	<,0001	<,0001
T*T	<,0001	<,0001	<,0001	<,0001
T*T*G	<,0001	<,0001	<,0001	<,0001

Silva (2011) trabalhando com medidas repetidas via modelos mistos na cultura da cana-de-açúcar nas usinas de Denusa, Goiasa/Série 95, Goiasa/Série 96 e Jalles Machado, observou que nos ensaios de Denusa e Goiasa /Série 95 as diferentes estruturas de covariâncias testadas não promoveram diferenças nos efeitos fixos do modelo (genótipos, cortes e genótipos x cortes). Porém, houve diferenças para os efeitos fixos nos ensaios de Goiasa/Série 96 e Jalles Machado. O autor reitera que a escolha inadequada da estrutura da matriz de variâncias e covariâncias pode resultar em inferências incorretas.

A tabela 15 apresenta os erros padrão das estimativas dos efeitos fixos dos modelos selecionados dentro de cada grupo. Pode-se observar uma redução expressiva nos erros padrão das estimativas dos modelos comparados. Comparando-se o melhor modelo do grupo U com o melhor modelo do grupo F, observa-se uma redução de 86,11% para o erro padrão do intercepto. Para a estimativa dos genótipos houve uma redução de 86,07% no erro padrão. Esses resultados confirmam a melhor qualidade de ajuste indicada pelos critérios de informação.

Tabela 15. Erros padrão das estimativas dos efeitos fixos dos modelos selecionados.

Efeito	Modelos			
	U	E	N	F
Intercepto	3,96	3,96	0,85	0,55
G	5,60	5,60	1,21	0,78
T	0,45	0,45	0,19	0,12
T*G	0,64	0,64	0,27	0,17
T*T	0,01	0,01	0,01	0,01
T*T*G	0,02	0,02	0,01	0,01

Na figura 3 encontra-se a representação gráfica das curvas de crescimento ajustadas para cada genótipo. Pode-se observar que os genótipos UENF 14-3-3 e FELTRIN (cultivar comercial) foram os que apresentaram o maior crescimento médio e os genótipos UENF 1445 (Progenitor UENF 19), UENF 9-24-2 e UENF 14-6-3 o menor. Além disso, como indicado na análise exploratória pode-se observar um crescimento inicial bastante semelhante entre os genótipos. As curvas polinomiais de segundo grau para cada genótipo são dadas abaixo:

$$\text{UENF 1445, } \hat{y} = 7,4318 - 0,6497t + 0,0888t^2$$

$$\text{FELTRIN, } \hat{y} = 13,6989 - 2,3700t + 0,2174t^2$$

$$\text{TOP SEED BLUE LINE, } \hat{y} = 13,8358 - 1,9514t + 0,1812t^2$$

$$\text{UENF 7-3-1, } \hat{y} = 10,9889 - 1,4838t + 0,14131 t^2$$

$$\text{UENF 7-4-1, } \hat{y} = 10,2391 - 1,732t + 0,1715 t^2$$

$$\text{UENF 7-5-1, } \hat{y} = 12,8157 - 2,053t + 0,18786 t^2$$

$$\text{UENF 7-6-1, } \hat{y} = 12,2741 - 2,2735t + 0,1887 t^2$$

$$\text{UENF 7-9-1, } \hat{y} = 9,3658 - 1,12205t + 0,122105 t^2$$

$$\text{UENF 7-10-1, } \hat{y} = 11,6161 - 1,5931t + 0,1471 t^2$$

$$\text{UENF 7-12-1, } \hat{y} = 8,5742 - 1,18633t + 0,120409 t^2$$

$$\text{UENF 7-14-1, } \hat{y} = 11,1202 - 1,7993t + 0,1687 t^2$$

$$\text{UENF 7-20-1, } \hat{y} = 13,6271 - 1,9232t + 0,17683 t^2$$

$$\text{UENF 9-24-2, } \hat{y} = 7,2659 - 0,6824t + 0,09682 t^2$$

$$\text{UENF 14-3-3, } \hat{y} = 15,0015 - 2,6547t + 0,2303 t^2$$

$$\text{UENF 14-4-3, } \hat{y} = 8,5661 - 1,3786t + 0,126175 t^2$$

$$\text{UENF 14-6-3, } \hat{y} = 7,2776 - 0,8277t + 0,10306 t^2$$

$$\text{UENF 15-23-4, } \hat{y} = 9,5155 - 1,1227t + 0,1193 t^2$$

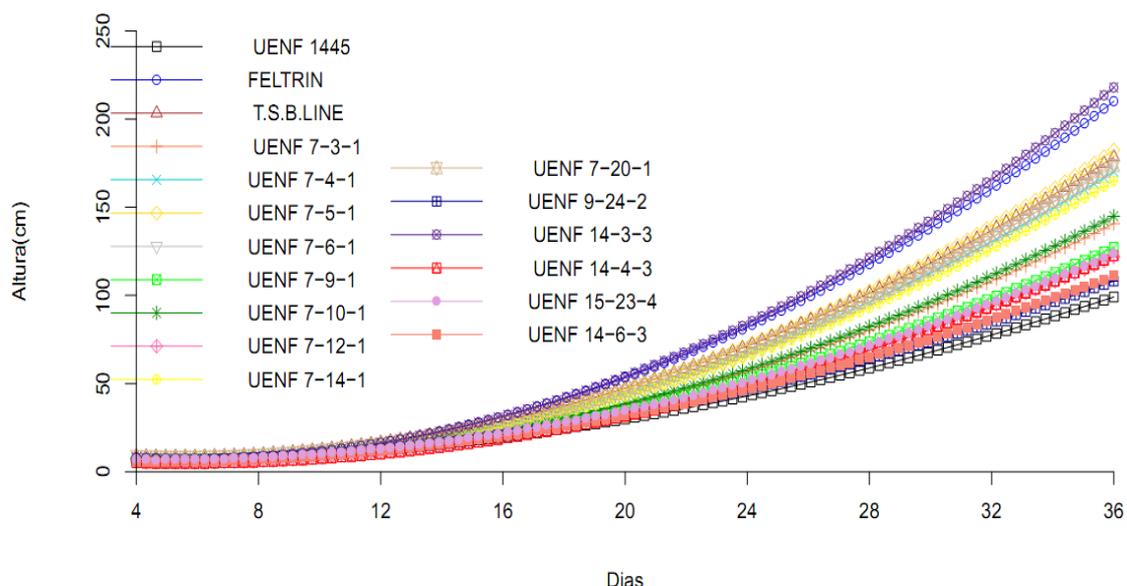


Figura 3: Curvas de crescimento ajustadas via modelos mistos para a variável altura de plantas dos 17 genótipos de feijão vagem ao longo do tempo.

Em trabalho de avaliação de linhagens melhoradas de feijão vagem em Bom Jesus do Itabopoana e estudando os mesmos genótipos considerados no presente trabalho Araujo (2011) não observou diferença entre os genótipos avaliados, apesar da variabilidade constatada. O autor destacou o genótipo 1445 como o que apresentou o maior crescimento. Resultado que discorda do observado no presente trabalho. Vale ressaltar, que no presente trabalho as avaliações não compreenderam todo o ciclo da cultura.

O crescimento dos genótipos estudados variou de 99,13 a 217,90 cm aos 36 dias. Avaliações realizadas por outros autores considerando todo o ciclo da cultura revelaram alturas variando de 0,79 a 1,83 cm (Abreu et al., 2004), de 1,47 a 2,49 cm (Francelino et al., 2011) e de 225 a 312 cm (Araujo, 2011).

5. RESUMOS E CONCLUSÕES

Na pesquisa científica dados longitudinais são bastante comuns na prática, sendo utilizados por pesquisadores de diversas áreas e constituem um caso especial dos estudos de medidas repetidas, termo que se refere a medidas feitas na mesma unidade experimental em mais de uma ocasião. Este tipo de dado requer um cuidado especial, pois as pressuposições necessárias para a execução da análise de variância podem não ser atendidas.

Dados longitudinais tendem a apresentar correlação entre as medidas feitas na mesma unidade experimental, e este fato deve ser levado em consideração na análise estatística para que as inferências possam ser válidas. Diversas são as análises propostas para a análise de dados longitudinais. Entretanto, nem todas elas se adéquam a um determinado conjunto de dados. No presente trabalho, pôde-se observar que a utilização da análise univariada de perfil não foi uma alternativa correta, visto que a condição H-F, que implica em variâncias e correlações homogêneas não foi atendida. Além disso, a análise univariada aproximada, realizada mediante a utilização de correções dos graus de liberdade também não se demonstrou como boa alternativa, visto que os resultados obtidos não se diferenciaram dos da análise univariada de perfil.

A análise multivariada é uma alternativa natural para o estudo de dados longitudinais devido ao fato de utilizar uma estrutura de covariância bastante geral e que não apresenta restrições quanto às variâncias, covariâncias e correlações. No presente trabalho, apesar do modelo multivariado considerar uma estrutura de

covariância sem restrição, ou seja, não estruturada as conclusões para os efeitos fixos não se alteraram comparando-se à análise univariada de perfil.

Por meio dos modelos mistos e dos diversos modelos gerados pode-se observar que a inclusão de efeitos aleatórios no modelo não melhorou a qualidade de ajuste. O modelo mais adequado dentre todos os testados foi formado por todos os efeitos fixos propostos, porém sem a variação entre – indivíduos, ou seja, sem os efeitos de intercepto, tempo e tempo². Para a modelagem dos erros a melhor estrutura de covariância foi a do tipo sem estrutura (UN) e a pior a estrutura componente de variâncias (VC). Além disso, pôde-se observar que os diferentes modelos selecionados não influenciaram a conclusão sobre os efeitos fixos. Entretanto, afetaram os erros padrão das estimativas dos efeitos fixos do modelo. Os resultados alcançados no presente trabalho reforçam a necessidade de se considerar as correlações existentes em dados longitudinais para que se tenha uma melhor qualidade de ajuste.

6.0 REFERÊNCIAS BIBLIOGRÁFICAS

- Abreu, F. B. (2001) *Aplicação de técnicas de análises multivariada em acessos de feijão-vagem (Phaseolus vulgaris L.) de crescimento indeterminado do banco de germoplasma da UENF*. Dissertação (Mestrado em Produção Vegetal). Campos dos Goytacazes – RJ, Universidade Estadual do Norte Fluminense Darcy Ribeiro - UENF, 69p.
- Abreu, F. B., Rodrigues, R., Leal, N. R., Júnior, A.T.A., Silva, D. J. H. (2000) Caracterização preliminar de acessos de feijão-vagem e feijão de corda do banco de germoplasma da UENF. *Horticultura Brasileira*. Brasília, 18: 741-742.
- Abreu, F.B., Leal, N.R., Amaral Júnior, A.T., Silva, D.J.H (2004) Divergência genética entre acessos de feijão-vagem de crescimento indeterminado. *Horticultura Brasileira*, 22 (3): 547-552.
- Agranonik, M. (2009) Equações de estimação generalizadas (GEE): Aplicação em estudo sobre mortalidade neonatal em gemelares de Porto Alegre, RS (1995-2007). Dissertação (Mestrado em Epidemiologia). Porto Alegre – RS, Universidade Federal do Rio Grande do Sul – UFRS, 110p.

- Aidar, H., Silva, S.C., Kluthcouski, J., Thung, M. (2002) Sistema de produção do feijoeiro comum em várzeas tropicais: época de plantio. Santo Antonio de Goiás EMBRAPA, 2002. (circular técnica, 55). Disponível em: <<http://www.cnpaf.embrapa.br> CNPAF, /publicação/circular técnica/ct_55)index.htm em 18/12/2013.
- Aldrighi, C. B. (2000) *Produção de feijão-vagem em duas épocas de cultivo em ambiente protegido com adubação orgânica*. Dissertação (Mestrado) – Pelotas – RGS, Faculdade de Agronomia, Universidade Federal de Pelotas - UFPel, 62p.
- Almeida, S.N.C (2011) *Avaliação de características morfológicas e agronômicas de linhagens de feijão de vagem em Bom Jesus do Itabopoana-RJ, com potencial de recomendação*. Dissertação (Mestrado em Produção Vegetal). Campos dos Goytacazes – RJ, Universidade Estadual do Norte Fluminense Darcy Ribeiro - UENF, 61p.
- Alves, E.U. (1999) *Produção e qualidade de sementes de feijão-vagem (Phaseolus vulgaris L.) em função de fontes e doses de matéria orgânica*. Dissertação de Mestrado - Centro de Ciências Agrárias, Universidade Federal da Paraíba - UFPB, 109p.
- Alves, U.E., OLiveira, A.D., Bruno, L.R.A., Araújo, E., Silva, J.A.L., Gonçalves, E.P., Costa, C.C (2000) Produção de sementes de feijão-vagem em função de fontes e doses de matéria orgânica. *Horticultura Brasileira*, 18 (3): 215-221.
- Amador, J.P. (2010) *Modelos mistos no ajuste de curvas de crescimento de Ricinus communis L.* Tese (Doutorado em Agronomia) – Universidade Federal de Santa Maria, 2010,139p.
- Amaro, G.B., Ramalho, M.A.P., ABreu, A.F.B., Silva, F.B. (2007) Phenotypic recurrent selection in the common bean (*Phaseolus vulgaris* L.) with carioca-type grains for resistance to the fungi *Phaeoisariopsis griseola*. *Genetics and Molecular Biology*, 30(3): 584-588.

- Aparicio, P.D., Ferreira, R.L.C., Silva, J.A.A., Rosa, A.C., Aparicio, W.C.S. (2010) Controle da matocompetição em plantios de dois clones de *Eucalyptus* × *urograndis* no Amapá. *Ciência Florestal*, Santa Maria, v. 20, n. 3, p. 381-390.
- Araújo, A. M. (2007) *Modelos não lineares para dados de contagem longitudinais*. Tese (Doutorado em Agronomia) – Escola Superior de Agricultura “Luiz de Queiroz” – Universidade de São Paulo, 2007, 197p.
- Araujo, J.S., Oliveira, A.P., Silva, J.A.L., Ramalho, C.I., Neto, F.L.C. (2001) Rendimento do feijão-vagem cultivado com esterco suíno e adubação mineral. *Revista Ceres*, 48(278): 501- 510.
- Araujo, L.C. (2011) *Avaliação de linhagens melhoradas de feijão de vagem, em Bom Jesus do Itabopoana, RJ*. Dissertação (Mestre em Produção Vegetal) – Universidade Estadual do Norte Fluminense Darcy Ribeiro, 2011, 59p.
- Athanázio, J. C. (1993) Adubação de feijão-vagem. In: Ferreira. M.E., Castellane. P.D., Cruz. M.C.P.da. *Nutrição e adubação de hortaliças*. Associação Brasileira para Pesquisa da Potassa e do Fosfato. 487p.
- Atkin, J.D. (1972) Natures of the stringy pod rogue of snap beans (*Phaseolus vulgaris*). *Search Agricultural*, 2(9): 1-3.
- Aubin, E. C. Q. (1984) *Análise de Experimentos com Medidas Repetidas*. Dissertação (Mestrado em Estatística) - Instituto de Matemática e Estatística, Universidade de São Paulo – IME - USP, 164p.
- Barbosa, M.L., Rezende, M.R.R., Costa, H.S.C., Maluf, W.R. (2001) *A cultura do feijão-vagem*. Boletim Técnico de Hortaliças. Lavras - MG, 1(65).
- Barbosa, M. (2009) *Uma abordagem para análise de dados com medidas repetidas utilizando modelos lineares mistos*. Dissertação (Estatística e Experimentação Agronômica) Piracicaba-SP, Escola Superior de Agricultura “Luiz de Queiroz”, 2009, 118p.

- Basso, J.M., Takahashi, H.W., Fregonezi, G.A.F. (2012) Cultivo de feijão-vagem em vaso em diferentes substratos e condutividades elétricas da solução nutritiva. *Semina*, 33(6): 2149-2162.
- Bates, D., Pinheiro, J.C. (2000) lmer for SAS PROC MIXED users. [s.l.: s.n.]. Disponível em: <<http://biostat.hitchcock.org/FacultyandSta@/OnlineManuals/PDF%20Files/lme%20sas.pdf>>. Acesso em: 10 jan. 2008.
- Bernardes, M.S. (1987) Fotossíntese no dossel das plantas cultivadas. In: Castro, P. R. G., Ferreira, S. O., Yamada, T. (Ed.). *Ecofisiologia da produção agrícola*. Piracicaba: Potafos, p. 13-48.
- Böhmer, C.R.K. (2008) *Caracterização do microclima de ambiente protegido cultivado com feijão-vagem*. Tese (Doutorado em Ciências), Pelotas – RGS, Universidade Federal de Pelotas – UFPel, 113p.
- Box, G. E. P. (1954a) Some theorems on quadratic forms applied to the study of analysis of variance problems I, II. *Ann. Math. Stat., Ann Arbor*, 25:290-302.
- Box, G.E.P. (1954b) Some theorems on quadratic forms applied in the study of analysis of variance problems II. Effects of inequality of variance and of correlation between errors in the two-way classification. *Annals of the Mathematical Statistics*, New York, 25:484-498.
- Camarinha Filho, J.A. (2002) *Modelos lineares Mistos: Estruturas de Matrizes de variância e covariância e seleção de modelo*. Tese (Doutorado em Agronomia) – Escola Superior de agricultura “Luis de Queiros”- Universidade de São Paulo, 74p.
- Carnicelli, J.H., Pereira, P.R.G., Fontes, P.C.R., Camargos, M.I. (2000) Índices de nitrogênio na planta relacionados com a produção comercial de cenoura. *Horticultura Brasileira*, 18:808-810, Suplemento.

- Carrijo, O. A., Silva, W. L. C., Marouelli, W. A., Silva, H. R. (1999) Tendências e desafios da fertirrigação no Brasil. In: Folegatti, M. V. *Fertirrigação: citrus, flores e hortaliças*. Guaíba: Agropecuária, p.155-169.
- Castellane, P.D., Vieira, R.D., Carvalho, N.M. (1998) *Feijão-vagem (Phaseolus vulgaris L.): cultivo e produção de sementes*. Jaboticabal: FCAV-UNESP, 60p.
- Castro, S.M.de.J. *A metodologia de análise de dados longitudinais*. (1997) Monografia (Bacharel em Estatística) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 119p.
- Castro, S. M. de J. (1999) *A metodologia de análise de dados longitudinais em plantas forrageiras no contexto de estruturas de covariância*. Dissertação (Mestrado em Zootecnia) – Universidade Federal do Rio Grande do Sul, 156p.
- Central de Abastecimento do Estado do Rio de Janeiro (2012)
- Cestari, A.A., Costa, S.C., Minho, A.P. (2012) Análise de dados longitudinais em experimentação animal. *Semina: Ciências Agrárias*, Londrina, 33(4):1565-1580.
- Consultores Estatísticos (2003) Diagnóstico dos processos de produção de hortaliças, Balsas – MA. CONSTAT. Disponível em [http://201.2.114.147/bds/BDS.nsf/F916B7DCAF02DB6403256F1E00511768/\\$File/NT0009D5FE.pdf](http://201.2.114.147/bds/BDS.nsf/F916B7DCAF02DB6403256F1E00511768/$File/NT0009D5FE.pdf) em 18/12/2013.
- Costa, S.C. (2003) *Modelos lineares generalizados mistos para dados longitudinais*. Tese (Doutorado em Estatística e Experimentação Agrônômica) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 110p.
- Crowder, M.J & Hand, D.J. (1990) *Analysis of repeated measures*. London: Chapman & Hall, 257p.

- Oca, G.M. (1987) Mejoramiento genético de la habichuela en el CIAT y resultados de viveros internacionales. In: Davis. J. & Jassem. W. (Ed.) El Mejoramiento genético de la habichuela in America Latina: memorias de un taller. Cali. Colômbia. CIAT. p.60-72 (Documentos de trabajo. 30). Estado de Minas. Viçosa: Imprensa Universitária, 569p. FAO (2010) http://www.fao.org/index_en.htm. Página mantida pela FAO.
- Fernandez, G.C.J. (1991) Repeated measure analysis of line-source sprinkler experiments. *HortScience*, v.26, p.339-342.
- Ferreira, W. L & Morais, A. R. (2013) Análise da influencia do café no ganho de peso de animais (ratos) por meio do modelo misto. *Rev. Bras. Biom.*, São Paulo, 31(4):485-500.
- Ferreira, W.L. (2012) *Análise de dados com medidas repetidas em experimento de café*. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) – Universidade Federal de Lavras, 108p.
- Filgueira, F.A.R. (2012) *Novo Manual de Olericultura: agrotecnologia moderna na produção e comercialização de hortaliças*. Viçosa: Universidade Federal de Viçosa – UFV, 421p.
- Filgueira, F.A.R. (2000) *Novo Manual de Olericultura: agrotecnologia moderna na produção e comercialização de hortaliças*. Viçosa: Universidade Federal de Viçosa – UFV, 402 p.
- Filgueira, F.A.R. (2008) *Novo Manual de Olericultura: Agrotecnologia moderna na produção e comercialização de hortaliças*. Viçosa: Universidade Federal de Viçosa – UFV, 421p.
- Fisher, R. (1918) The correlation between relatives on the supposition of Mendelian inheritance. *Royal Society of Edinburgh Transactions*, Edinburgh, 52: 399-433.

- Floriano, E.P., Müller, I., Finger, C.A.G., Schneider, R. (2006) Ajuste e seleção de modelos tradicionais para serie temporal de dados de altura de arvores. *Ciência Florestal*, Santa Maria, 16(2): 177-199.
- Francelino, F.M.A., Gravina, G.A., Manhaes, C.M.C., Cardoso, P.M.R., Araujo, L.C. (2011) Avaliação de linhagens de feijão de vagem para as regiões Norte e Noroeste Fluminense. *Rev. Ciênc. Agron.*, 42(2):554-562.
- Freitas, E.G. (2007) *Análise de dados longitudinais em experimentos com cana de açúcar*. Dissertação (Mestrado em Agronomia) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 74p.
- Freitas, E.G., Barbin, D., Barbosa, G.V.S., Carneiro, M. S., Bassinello, A.I. (2008) Modelo univariado aplicado à dados longitudinais de cana de açúcar. *Revista Brasileira de Biometria*, São Paulo, 26(2):93-106.
- Garcia, C.H. (2004) *Aplicações de equações de modelo mistos em testes clonais de Eucalyptus spp.* Dissertação (Mestre em Agronomia) – Piracicaba - SP, Escola Superior de Agricultura "Luiz de Queiroz"- Universidade de São Paulo – ESALQ – USP, 74p.
- Geisser, S., Greenhouse, S.W. (1958) An extension of Box's results on the use of the F distribution in multivariate analysis. *Annals of the Mathematical Statistics*, Ann Arbor, 29:855-891.
- Gill, J.L. (1986) Design and analysis of experiments in the animal and medical sciences. Ames: The Yowa State University Press, 301p.
- Guedes, R. E. (2003) *Manejo orgânico de cultivares arbustivas de feijão-vagem no Estado do Rio de Janeiro*. Dissertação (Mestrado em Fitotecnia), Seropédica – RJ, Universidade Federal Rural do Rio de Janeiro - UFRRJ, 98p.

- Gepts, P., Debouck, D.G. Origin, domestication, and evolution of the common bean (*Phaseolus vulgaris*). In: SCHOONHOVEN, A. van; VOYSEST, O. (Ed.). Common beans: research for crop improvement. Cali: CIAT, 1991. p.7-53.
- Guimarães, C.M., Stone, L.F., Neves, P. de C.F. (2008) Eficiência produtiva de cultivares de arroz com divergência fenotípica. *Revista Brasileira de Engenharia Agrícola e Ambiental*, 12:465-470.
- Guimarães, C. M. (1988) Efeitos fisiológicos do estresse hídrico. In: Zimmermann, M. J. de O., Rocha, M., Yamada, T. *Cultura do feijoeiro: fatores que afetam a produtividade*. Piracicaba: Potafos, p.157-174.
- Hartley, H.O., Rao, J. N. K. (1967) Maximum likelihood estimation for the mixed analysis of variance model. *Biometrika*, 54(1/2):93-108.
- Hauser, L., Vigo, A., Kieling, C., Jardim, L.B., Camey, S.A., Leotti, V.B. (2009) Curva de crescimento usando modelo misto: Uma aplicação na progressão da doença de Machado-Joseph. *Revista do Hospital de Clínicas de Porto Alegre*, 29(1): 05-17.
- Henderson, C.R. (1975) The best linear unbiased estimation and prediction under a selection model. *Biometrics*, 31:423-449.
- Henderson, C.R. (1953) Estimation of variance and variance components. *Biometrics*, 9:226-252.
- Heldwein, A.B., Streck, N.A., Sturza, V.S., Loose, L.H., Zanon, A.J., Toebe, M., Souza, A.T., Peters, M.B., Karlec, F. (2010) Plastocrono e rendimento de feijão-vagem cultivado sob ambiente protegido e no ambiente externo em semeadura tardia no outono. *Ciência Rural*, 40(4): 768-773.
- Henry, G & Jansen, W. (1992) *Snap Beans in the developing world*. Calí. Centro internacional de agricultura tropical. 366p.

- Huynh, H & Feldt, L.S. (1976) Estimation of the box correction for degrees of freedom from sample data in the randomized block and split-plot designs. *Journal of Educational and Behavioral Statistics*, Boston, 1(1):69-82.
- Hunt, R. (1990) Basic growth analysis. London: Unwin Hyman, 112 p
- Instituto Brasileiro de Geografia e Estatística (2006) Disponível em: <http://www.sidra.ibge.gov.br/> em 18/12/2013.
- Kaplan, L. (1981) What is the origin of the common bean. *Economic Botany*, 35(2):40-257.
- Ker, H. (2010) Visual-Graphical Methods for Exploring Longitudinal Data. *World Academy of Science, Engineering and Technology*, 62:445-452.
- Krause, W. (2008) *Fontes de resistência, métodos de inoculação e Capacidade de combinação para a resistência à Murcha-de-curtobacterium em feijão-vagem*. Tese (Doutorado em Genética e Melhoramento de Plantas), Campos dos Goytacazes – RJ, Universidade Estadual do Norte Fluminense Darcy Ribeiro - UENF, 108p.
- Kurek, A.J., Carvalho, F.I.F., Assmann, I.C., Marchioro, V.S., Cruz, P.J. (2001) Análise de trilha como critério de seleção indireta para rendimento de grãos em feijão. *Revista Brasileira de Agrociência*, 7(1):29-32.
- Laird, N.M & Ware, J.H. (1982) Random effects models for longitudinal data. *Biometrics*, Washington, 38:963-974.
- Leal, N.R. (1983) Introdução e adaptação da cultivar de feijão-vagem de porte determinado "Cascade". In: Congresso brasileiro de olericultura, 23, Rio de Janeiro, 1983. Anais... Rio de Janeiro, SOB, 1983. p.145.
- Liang, K-Y & Zeger, S.L. (1986) Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1):13-22.

- Lima, C.G. (1996) *Análise de dados longitudinais provenientes de experimentos em blocos casualizados*. Tese (Doutorado em Estatística e Experimentação Agronômica) - Piracicaba - SP, Escola Superior de Agricultura "Luiz de Queiroz"- Universidade de São Paulo – ESALQ – USP, 119p.
- Lima, C.G. (1988) *Análise de curvas de crescimento de aves – um enfoque multivariado*. Dissertação (Mestrado em Agronomia) – Escola Superior de Agricultura “Luiz de Queiroz” – Universidade de São Paulo, 69p.
- Littell, R.C., Henry, P.R., Ammerman, C.B (1998) Statistical analysis of repeated measures data using SAS procedures. *Journal Animal Science*, 76:1216-1231.
- Littell, R.C., Milliken, G.A., Stroup, W.W., Wolfinger, R.D., Schabenberger, O. (2006) *SAS mixed models*. 2 ed. Cary: SAS Institute, 814 p.
- Lopes, N. R., Oliva, M. A., Cardoso, M. J., Gomes, M. M. S., Souza, V. F. (1987) Crescimento, morfologia, partição de assimilados e produção de matéria seca de feijão (*Phaseolus Vulgaris* L.) submetido a três densidades de fluxo radiante e dois regimes hídricos. *Revista Ceres*, 34(132):110-124.
- Macedo, J. O. (2007) *Modelos de regressão aleatória usando como bases as funções polinomiais de Legendre, de Jacobi modificadas e trigonométricas, com uma aplicação na análise genética dos pesos de bovinos da raça nelore*. Tese (Doutorado em Agronomia) - Piracicaba - SP, Escola Superior de Agricultura "Luiz de Queiroz"- Universidade de São Paulo – ESALQ – USP, 100p.
- Malheiros, E.B. (2004) Precisão de teste F univariados usados em experimentos com medidas repetidas no tempo, quando a condição de esfericidade da matriz de covariâncias não é verificada. *Rev. Mat. Estat.*, São Paulo, 22(2):23-29.

- Mariguele, K. H., Resende, M.D.V., Viana, J.M.S., Silva, F.F., Silva, P.S.L., Knop, F.C. (2011) Métodos de análise de dados longitudinais para o melhoramento genético da pinha. *Pesq. agropec. bras.*, Brasília, 46(12):1657-1664.
- Mariguele, K.H., Maluf, W.R., Gomes, L.A.A., Lopes, M.J.C., Melo, O.D (2008) Controle genético da qualidade da vagem em cruzamento de feijão-vagem e feijão-comum. *Pesq. agropec. bras.*, 43(1):47-52.
- Maroto, J.V. (2000) *Horticultura herbacea especial*. Madri: Artes Gráficas Cuesta S.A., 611p.
- Mauchly, J. W. (1940) Significance test for sphericity of a normal n-variate distribution. *Annals of Mathematical Statistics*, Beachwood, 11(2): 204-209.
- Menezes Júnior, A.O., Simões, H.C., Souza, A.M., Takada, R.M. (2003) Flutuação populacional de insetos associados ao feijão-vagem cv. UEL-1 em quatro épocas de plantio e seu efeito sobre as características produtivas. *Semina: Ciências Agrárias*, 24(2): 283-288.
- Meredith, M.P & Stehman, S.V. (1991) Repeated measures experiments in forestry: focus on analysis of response curves. *Canadian Journal of Forest Research*, 21:957-965.
- Milliken, G.A & Johnson, D.E. (1992) Analysis of messy data: designed experiments. New York: Chapman & Hall, 1, 473p.
- Nóbrega, J.Q., Tantravahi, V.R.R., Beltrão, N. E. DE M., Fideles Filho, J. (2001) Análise de crescimento do feijoeiro submetido a quatro níveis de umidade do solo. *Revista Brasileira de Engenharia Agrícola e Ambiental*, 5(3): 437-443.
- Normas de Classificação (2012) Programa Brasileiro para a Modernização da Horticultura: NB42, 10(1): 1-6.

- Oliveira, A.P., Sobrinho, J. T., Souza, A.P. (2003) Característica e rendimento do feijão-vagem em função de doses e formas de aplicação de nitrogênio. *Ciência e Agrotecnologia*, 37(3): 714-720.
- Oliveira, A.P., Andrade, A.C., SObrinho, J.T., Peixoto, N. (2001) Avaliação de linhagens e cultivares de feijão-vagem de crescimento indeterminado, no município de Areia-PB. *Horticultura Brasileira*, 19(2): 159-162.
- Oliveira, A.P., Cardoso, M.O., BARbosa, L.J.N., Silva, J.E.L., Morais, M.S. (2005) Resposta do feijão-vagem a P₂O₅ em solo arenoso com baixo teor de fósforo. *Horticultura Brasileira*, 23(1):128-132.
- Oliveira, A.P., Silva, J.A., Alves, A.U., Dorneles, C.S.M., Alves, A.U., Oliveira, A.N.P., Cardoso, E.A., Silva, Cruz, I.S. (2007) Rendimento de feijão-vagem em função de doses de K₂O. *Horticultura Brasileira*, 25:29-33.
- Oliveira, G.N., De-Polli, H., Almeida, D.L., Guerra, J.G.M. (2006) Feijão-vagem semeado sobre cobertura viva perene de gramínea e leguminosa e em solo mobilizado, com adubação orgânica. *Pesq. agropec. bras*, 41(9):1361-1367.
- Omae, H., Kumar, A., Kashiwaba, K., Shono, M. (2007) Influence of temperature shift after flowering on dry matter partitioning in two cultivars of snap beans (*Phaseolus vulgaris*) that differ in heat tolerance. *Plant Production Science*, 10(1):14-19.
- Pascual, M.G.C.A. (2000) *Metodologia de dados espaciais aplicados a plantas forrageiras*. Tese (Doutorado em Zootecnia) Porto Alegre – RS, Faculdade de Agronomia - Universidade Federal do Rio Grande do Sul. – UFRS, 100p.
- Patterson, H. D., Thompson, R. (1971) Recovery of inter-block information when block sizes are unequal. *Biometrika*, 58:545-554.

- Peixoto, N., Silva, L.O., Thung, M.D.T., SANTOS, G. (1993) Produção de sementes de linhagens e cultivares arbustivas de feijão-vagem em Anápolis - GO. *Horticultura Brasileira*, 11(2): 151-152.
- Peixoto, N., Braz, L. T., Banzatto, D.A., Oliveira, A.P (2002) Adaptabilidade e estabilidade em feijão-vagem de crescimento indeterminado. *Horticultura Brasileira*, 20(4): 616-618.
- Peixoto, N., Moraes, E.A., Monteiro, J.D., Thung, M.D.T. (2001) Seleção de linhagens de feijão-vagem de crescimento indeterminado para cultivo no Estado de Goiás. *Horticultura Brasileira*, 19(1): 85-88.
- Peixoto, N., THung, M.D.T., Silva, L.O., Farias, J.G., Oliveira, E.B., Barbedo, A.S.C., Santos, G. (1997) *Avaliação de cultivares arbustivas de feijão-vagem, em diferentes ambientes do Estado de Goiás*. Boletim de Pesquisa 01 Emater-GO, 20p.
- Perri, S. H. V., Iemma, A. F. (1999) Procedure of software SAS r for the analysis of mixed models. *Scientia Agricola*, Piracicaba, 56(4).
- Pinheiro, J. C. (1994) Topics in mixed effects models. Thesis (PhD) - Madison – EUA, University of Wisconsin, 210 p
- Pinheiro, J.C.; Bates, D.M. (2000) Mixed-effects models in S and S-PLUS. New York: Springer-Verlang, 528 p.
- Pinto, C.M.F., Vleira, R.F., Vieira, C., Caldas, M. T. (2001a) Comportamento de cultivares de feijão-vagem anão em diferentes épocas de plantio na Zona da Mata de Minas Gerais. *Horticultura Brasileira*, 19, Suplemento CD-ROM.
- Pinto, C.M.F., Vieira, R.F., Vieira, C., Caldas, M.T. (2001) Idade de colheita do feijão-vagem anão cultivar Novirex. *Horticultura Brasileira*, 19(1): 163-167.

- Pires, L.P.M. (2013) *Seleção recorrente massal para porte ereto em feijão do tipo carioca*. Dissertação (Mestrado em Genética e Melhoramento de plantas), Lavras – MG, Universidade Federal de Lavras – UFLA. 81p.
- Portes, T.A. (1988) Ecofisiologia. In: Zimmermann, M. J. O., Rocha, M., Yamada, T. *Cultura do feijoeiro: fatores que afetam a produtividade*. Piracicaba: Potafós, p.125.
- Rao, C.R. (1970) Estimation of heterocedastic variances in linear models. *Journal of the American Statistical Association*, 65:161-172.
- Rao, C.R. (1971a) Estimation of variance and covariance components - MINQUE Theory. *Journal of Multivariate Analysis*, 1:257-275.
- Rao, C.R. (1971b) Minimum variance quadratic unbiased estimation of variance components. *Journal of Multivariate Analysis*, 1: 445-456.
- Rezende, D.M.L.C. et al. Comparações de técnicas de análises de experimentos utilizando medidas repetidas no tempo. *Ciência e Agrotecnologia*, v.23, n.4, p.927-937, 1999.
- Resende, M.D.V. *Matemática e estatística na análise de experimentos e no melhoramento genético*. Colombo: Embrapa Floresta, 2007. 362 p.
- Riboldi, J., Fernandez, D.W.X., Castro, S.M.J., Análise de observações simultâneas e medidas repetidas. In: Reunião anual da região brasileira da sociedade internacional de biometria, 41., 1996, São José do Rio Preto. Resumos... São José do Rio Preto: UNESP, 1996.
- Riboldi, J. (1999) Análise de medidas repetidas: Perspectivas e Tendências. In: Reunião anual da região brasileira da sociedade internacional de biometria. 44 e Simpósio de estatística aplicada à experimentação agrônômica. 8., Botucatu. UNESP. Minicurso. Botucatu: Rbras, Seagro. Minicurso, 92 p.

- Riboldi, J. (2007) Modelos mistos aplicados. reunião anual da região brasileira da sociedade internacional de biometria. 52. Simpósio de estatística aplicada à experimentação agrônômica. 12., Santa Maria. UFSM. Minicurso. Santa Maria: Rbras, Seagro. Minicurso, 68 p.
- Rocha, M. G. B., Pires, I. E., Xavier, A., Cruz, C. D., Rocha, R. B. (2004) Avaliação genética de progênies de meios-irmãos de *Eucalyptus urophylla* utilizando os procedimentos REML/BLUP e E(QM). *Ciência Florestal*, 16(4): 369-379.
- Rosário, M.F. (2003) *Emprego do conceito de medidas repetidas na avaliação do desempenho de genótipos de frangos de corte*. Dissertação (Mestrado em Genética e Melhoramento de Plantas) - Piracicaba - SP, Escola Superior de Agricultura "Luiz de Queiroz"- Universidade de São Paulo – ESALQ – USP, 66p.
- Santos, A., Savian, T.V., Muniz, J.A. (2013) Regressão não linear no desdobramento da interação em experimentos com parcela subdividida no tempo. *Rev. Bras. Biom.*, São Paulo, 31(1):379-396.
- Santos, G.M., OLiveira, A.P., SILVA, J.A.L., ALves, E.U., Costa, C.C. (2001) Características e rendimento de vagem do feijão-vagem em função de fontes e doses de matéria orgânica. *Horticultura Brasileira*, 19(1): 30-35.
- Scheiner S.M & Gurevitch J. (2001) *Design and Analysis of Ecological Experiments*. 2 ed. Oxford: Oxford University Press, Inc: 445p.
- Searle, S.R. (1987) *Linear models for unbalanced data*. New York: J John Wiley, 536p.
- Shaw, R.G. (1987) Maximum-likelihood approaches to quantitative genetics of natural populations. *Evolution*, 41:812-826.

- Silbernagel, M.J., Janssen, W., Davis, J.H.C., Oca, G.M. (1991) Snap bean production in the tropics: implications for genetic improvement. In: choowhoven, A., Voysest, O. (eds.) Common beans: Research for crop improvement. Wallingford: CAB international, 835-862.
- Silva, A.B. (2013) *Cruzamentos dialélicos para caracteres agronômicos na cultura de feijão-vagem*. Dissertação (Mestrado em Genética e Melhoramento de Plantas) Campos dos Goytacazes – RJ, Universidade Estadual do Norte Fluminense Darcy Ribeiro - UENF, 54p.
- Silva, E.C., Silva Filho, A.V., Alvarenga, M.A.R. (2001) Efeito residual da adubação efetuada no cultivo da batata sobre a produção do feijão-vagem. *Horticultura Brasileira*, 19(3):180-183.
- Silva, E.N. (2011) *Análise de medidas repetidas em ensaios varietais da cana-de-açúcar*. Dissertação (Mestre em Agronomia) – Universidade Federal de Goiás, 96p.
- Silva, R.A., Souza, J.C., Rodrigues, J.R.M., Carvalho, T.A.F., Pereira, A.B. (2012a) *Principais pragas do feijão-vagem no Sul de Minas Gerais*. EPAMIG. Circular Técnica, 173.
- Silva, W.G., Carvalho, J.A.C., Oliveira, E.C., Rezende, F.C., Lima Junior, J.A., Rios, G.F.A. (2012) Manejo de irrigação para o feijão de metro, nas fases vegetativa e produtiva, em ambiente protegido. *Rev. bras. eng. agríc. Ambient*, 16(9):978-984.
- Silveira, P.M. da & Stone, L. F. (2004) Irrigação. Informe Agropecuário, Belo Horizonte, 25(223): 74-82.
- Singer, J. da. M. (1977) *Análise de curvas de crescimento*. Dissertação (Mestrado em Estatística) Instituto de Matemática e Estatística -- Universidade de São Paulo, USP, 108p.

- Singer, J.M., Nobre, J.S., Rocha, F.M.M (2012) *Análise de dados longitudinais*. Caderno Didático (Versão preliminar). Departamento de Estatística - Universidade de São Paulo, 209p.
- Souza, C.M., Silva, J.A.A., Ferreira, R.L.C., Santos, E.S., Meunier, I.M.J., Ferraz.I. (2008) Avaliação do crescimento em altura de *Leucaena leucocephala* (Lam.) de Wit., no Agreste de Pernambuco. *Rev. Bras. Ciênc. Agrár.* Recife, 3(3):260-266.
- Souza, C.M.P. (2012) *Análise dialéctica para caracteres quantitativos e Qualitativos entre genótipos de feijão-vagem (Phaseolus vulgaris L.) de porte determinado*. Tese (Doutorado em Produção Vegetal). Campos dos Goytacazes – RJ, Universidade Estadual do Norte Fluminense Darcy Ribeiro - UENF, 103p.
- Souza, J.L.de & Escobedo, J.F. (1997) Balanço de radiação em cultivos de feijão-vagem com e sem cobertura de polietileno. *Pesq. Agropec. Bras*, 32(31):1-15.
- Statistical Analysis System Institute - SAS. The SAS-System for Windows. User's Guide. versão 9.1.3. Cary: Statistical Analysis System Institute, 2003.
- Suyama, E. (1995) Modelos de Efeitos Aleatórios para Dados Longitudinais. In: reunião anual da região brasileira da sociedade internacional de biometria. 40 e simpósio de estatística aplicada à experimentação agrônômica. Ribeirão Preto. USP. Minicurso. Ribeirão Preto: Rbras, Seagro. Minicurso, 94 p.
- Tavares Sobrinho, J. (2001) *Produção e qualidade de feijão-vagem (Phaseolus vulgaris L.) em função de doses e formas de aplicação de nitrogênio*. Dissertação (Mestrado) - Centro de Ciências Agrárias, Universidade Federal da Paraíba - UFPB, Areia. 100p.
- Timm, N.H. (1980) Multivariate analysis of variance of repeated measurements. Handbook of Statistics, v 1. P. R. Krishnaiah (Ed). North-Holland:New York. p. 41-87.

- Venezuela, M.K. (2003) *Modelos lineares generalizados para análise de dados com medidas repetidas*. Dissertação (Mestrado em Estatística) - Instituto de Matemática e Estatística, 95p.
- Verneque, R.S. (1994) *Procedimentos numéricos e estimação de componentes de variância em análise multivariada pelo método da máxima verossimilhança restrita - modelos mistos aplicados ao melhoramento animal*. Tese (Doutorado) - Piracicaba - SP, Escola Superior de Agricultura "Luiz de Queiroz"- Universidade de São Paulo – ESALQ – USP, 157p.
- Vidal, V.L., Junqueira, A.M.R., Peixoto, N., Moraes, E.A. (2007) Desempenho de feijão-vagem arbustivo, sob cultivo orgânico em duas épocas. *Horticultura Brasileira*, 25:10-14.
- Vieira, C. (2006). O feijoeiro comum: cultura, doenças e melhoramento. Viçosa: Universidade Rural do Estado de Minas Gerais. 486 p.
- Vivaldi, L. J. (1999) *Análise de experimentos com dados repetidos ao longo do tempo ou espaço*. Planaltina: Empresa Brasileira de Pesquisa Agropecuária, Embrapa Cerrados, 52 p.
- Von Ende, C.N. Repeated-measures analysis: growth and other time-dependent measures. In: SCHEINER, S.M.; GUREVITCH, J. (Ed.) *Design and analysis of ecological experiments*. New York: Chapman and Hall, 1993. chap.6, p. 113-137.
- Vonesh, E.F & Chinchilli, V.M. (1997) *Linear and nonlinear models for the analysis of repeated measurements*. New York: Marcel Dekker, 560 p.
- Wald, V.B. (2000) *A metodologia de modelos mistos não lineares aplicados a análise de dados longitudinais em plantas forrageiras*. Dissertação (Mestrado em Zootecnia) – Universidade Federal do Rio Grande do Sul, 106p. .

West, B. T., Welch, K. B., Galecki, A.T. (2007) *Linear mixed models: a practical guide using statistical software*. New York: Chapman & Hall, 339p.

Xavier, L. H & Santos, C. T. D. (2001) Acurácia do Modelo Univariado para Análise de Medidas Repetidas por Simulação Multidimensional. *Scientia Agricola*, 58(2):241-250.

Xavier, L.H. (2000) *Modelos univariados e multivariados para análise de medidas repetidas e verificação da acuraria do modelo univariado por meio de simulação*. Dissertação (Mestrado em Estatística e Experimentação Agronômica) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 91p.