



Universidade Estadual da Paraíba
Centro de Ciências e Tecnologia
Departamento de Estatística

Eder Cabral Nobrega

Estudo do crescimento do girassol através da análise de perfil

Campina Grande
Dezembro de 2014

Eder Cabral Nobrega

Estudo do crescimento do girassol através da análise de perfil

Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba em cumprimento às exigências legais para obtenção do título de bacharel em Estatística.

Orientador:
Dr. João Gil de Luna

Campina Grande
Dezembro de 2014

É expressamente proibida a comercialização deste documento, tanto na forma impressa como eletrônica. Sua reprodução total ou parcial é permitida exclusivamente para fins acadêmicos e científicos, desde que na reprodução figure a identificação do autor, título, instituição e ano da dissertação.

N754e Nobrega, Eder Cabral.

Estudo do crescimento do girassol através da análise de perfil
[manuscrito] / Eder Cabral Nóbrega. - 2014.

39 p. : il.

Digitado.

Trabalho de Conclusão de Curso (Graduação em Estatística) -
Universidade Estadual da Paraíba, Centro de Ciências e
Tecnologia, 2014.

"Orientação: Prof. Dr. João Gil de Luna, Departamento de
Estatística".

1. Estatística. 2. Medidas repetidas. 3. Análise multivariada.
4. Cultura do girassol. I. Título.

21. ed. CDD 519.53

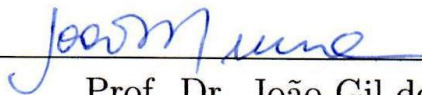
Eder Cabral Nobrega

Estudo do crescimento do girassol através da análise de perfil

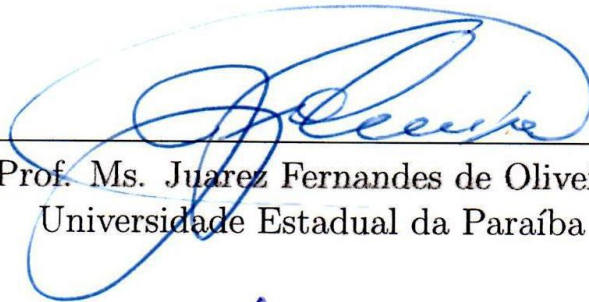
Trabalho de Conclusão de Curso apresentado ao curso de Bacharelado em Estatística do Departamento de Estatística do centro de Ciências e Tecnologia da Universidade Estadual da Paraíba em cumprimento às exigências legais para obtenção do título de bacharel em Estatística.

Aprovado em: 10 / dezembro / 2014

Banca Examinador:



Prof. Dr. João Gil de Luna
Orientador



Prof. Ms. Juarez Fernandes de Oliveira
Universidade Estadual da Paraíba



Prof. Dr. Tiago Almeida de Oliveira
Universidade Estadual da Paraíba

Agradecimentos

A Deus, fonte de vida, por iluminar meu caminho e sempre me dar forças para seguir em frente.

A minha mãe e meu pai (in memoriam) pela confiança em mim e por terem feito de suas vidas uma completa doação para que eu atingisse meus objetivos.

Aos meus irmãos por toda a ajuda durante o curso e em tantos momentos em minha vida.

Ao meu orientador, João Gil de Luna, por sua disponibilidade, dedicação e pelas discussões enriquecedoras que possibilitaram a conclusão do presente trabalho.

A Isaedja, pelas palavras de incentivo e pela confiança em mim.

Aos meus amigos de curso por todo o apoio durante essa jornada.

Enfim, a todos que de alguma maneira contribuíram para a execução desse trabalho.

Resumo

O aumento da demanda de grãos de girassol para extração do óleo para alimentação e produção de biocombustível se reflete no mercado de sementes, pois nos últimos anos a exigência por sementes de qualidade tem sido cada vez maior. Dessa forma, o objetivo deste trabalho foi avaliar as variáveis que se relacionam com a biomassa do girassol e assim verificar o crescimento da planta por meio do diâmetro do caule através da análise multivariada com medidas repetidas. O delineamento experimental de blocos casualizados teve quatro repetições em esquema de parcelas subdivididas no tempo, foi utilizado o esterco de caprinos como matéria orgânica em seis níveis de quantidade por litro na cova da planta. Após o início do experimento foram observadas as medidas do diâmetro do caule das plantas durante cinco avaliações no período de 30, 45, 60, 70 e 80 dias respectivamente. Os resultados apresentados mostraram que as medidas ao longo do tempo são correlacionadas entre si, quanto as hipóteses testadas pode-se concluir que existem evidências de que o período de dias nas avaliações e a quantidade de matéria orgânica utilizada têm influencia significativa no diâmetro do caule do girassol. Assim, a análise multivariada no modelo de medidas repetidas ajustou-se muito bem aos dados, permitindo uma efetiva avaliação sobre o crescimento da planta de girassol. Além disso, este trabalho pode verificar a importância das técnicas estatísticas para análise de dados avaliados ao longo do tempo.

Palavras-chave: Medidas repetidas, análise multivariada, cultura do girassol

Abstract

The increased demand for sunflower seeds to extract the oil for food and biofuel production is reflected in the seed market, because in recent years the demand for quality seeds has been increasing. Thus, the objective of this study was to evaluate the variables that relate to the biomass of sunflower and thus check the growth of the plant by the stem diameter by multivariate analysis with repeated measures. The experimental randomized block design was replicated four times in the time split plot was used manure goats as organic matter in six number of levels per liter in the pit of the plant. After the start of the experiment showed the stem diameter measurements of plants for five reviews within 30, 45, 60, 70 and 80 days respectively. The results showed that measures over time are correlated, as tested hypotheses can be concluded that there is evidence that the period of days on the reviews and the amount of organic matter used have significant influence on sunflower stem diameter. Thus, the multivariate analysis in the model of repeated measures set well to the data, allowing an effective review of the sunflower plant growth. In Addition addition, this work can verify the importance of statistical techniques for data analysis evaluated over time.

Key-words: Repeated measures multivariate analysis , sunflower cultivation

Sumário

Lista de Figuras

Lista de Tabelas

1	Introdução	p. 10
2	Referencial teórico	p. 13
2.1	Aspectos gerais sobre o girassol	p. 13
2.2	Fundamentos estatísticos	p. 14
2.2.1	Medidas repetidas	p. 14
2.2.1.1	Teste de esfericidade de Mauchly	p. 17
2.2.1.2	Análise multivariada de perfis	p. 19
3	Material e Métodos	p. 24
3.1	Caracterização dos dados	p. 24
3.2	Métodos utilizados	p. 25
3.2.1	Hipóteses testadas	p. 27
3.2.2	Procedimento da análise	p. 28
4	Resultados e Discussão	p. 29
5	Conclusões	p. 35
	Referências	p. 36

Lista de Figuras

1	Gráfico qq-plot de cada época de avaliação ao longo do ciclo do crescimento das plantas	p. 30
2	Perfil do diâmetro médio do caule do girassol (mm) em relação as épocas de avaliação em cada tratamento	p. 33
3	Boxplot dos dados originais	p. 34
4	QQ-plot dos dados originais	p. 34

Lista de Tabelas

1	Relação das medidas do diâmetro do caule do girassol (mm) em cada época de avaliação	p. 25
2	Teste Shapiro-Wilk (W) para cada época de avaliação	p. 31
3	Variância-Covariância de cada época de avaliação	p. 31
4	Correlação de cada época avaliação	p. 31
5	Resultado da análise multivariada, testando o efeito da época de avaliação.	p. 32
6	Resultado da análise multivariada, testando o efeito da interação da época de avaliação \times níveis de matéria orgânica.	p. 32
7	Soma de quadrados (S.q.), graus de liberdade (g.l.), Média quadrática (M.q), valor F e p -valor de contrastes formados entre as épocas de avaliação.	p. 33

1 Introdução

No Brasil, o cultivo do girassol vem crescendo gradativamente devido a demanda por grãos destinados à extração do óleo para utilização como matéria prima para a produção de biocombustível, cosméticos e para alimentação. A exploração do girassol para a produção de biodiesel se dá pelo seu alto teor de óleo (em torno de 40%), sendo esta exploração impulsionada pela criação do Programa Nacional de Biodiesel que traz como metas a inserção da agricultura familiar na produção de agroenergia (SILVA et al. 2008). Dessa forma, a cultura do girassol representa uma alternativa de grande importância por conciliar renda à atividade agrícola e ser fonte de proteínas de alto valor biológico para alimentação humana e animal, além de se constituir em uma das oleaginosas utilizadas para a produção de biodiesel.

Souza (2010), salienta que uma das estratégias para obter maiores produções, encontra-se na interceptação da energia radiante pela cobertura completa do solo, mediante a manipulação da densidade de semeadura, já que o crescimento de um vegetal depende, em termos gerais, do desenvolvimento e da expansão celular. Nesse caso, a análise de crescimento se apresenta como uma técnica viável para se conhecer as bases fisiológicas da produção tornando evidentes as influências exercidas pelas interações ambientais, genéticas e agronômicas. Esta técnica descreve as condições morfofisiológicas da planta em função do tempo.

A análise de crescimento é um meio acessível e preciso de se inferir a contribuição de diferentes processos fisiológicos sobre o comportamento vegetal (BENINCASA, 2003). Entretanto, as variáveis associadas ao crescimento das plantas de girassol, que são avaliadas periodicamente ao longo do ciclo vegetativo, influenciam fortemente nas práticas agrícolas necessárias ao desenvolvimento desta atividade. Por exemplo, a altura das plantas, o diâmetro do caule, a quantidade de folhas, o tamanho dos aquênios e etc. Dessa forma, impõe-se ao produtor o planejamento e adaptação das máquinas para a realização dos tratos culturais e colheita das sementes que é o seu principal produto, como também, o destino a ser dado para a biomassa produzida com a atividade: se vai ser incorporado ao

solo ou se vai ser aproveitado como ração animal, ou se vai ser aproveitado na produção de energia.

É concebível que as variáveis de crescimento do girassol estejam relacionadas entre si, como também, estão associadas a idade da planta e aos efeitos dos tratamentos a que as plantas são submetidas. Sendo assim, um tratamento estatístico desses dados, demanda técnicas que a priori, leve em consideração estas condições.

A análise multivariada é uma técnica estatística que processa as informações de modo a simplificar a estrutura dos dados e a sintetizar as informações, quando o número de variáveis envolvidas é muito grande, promovendo assim uma melhor interpretação entre as variáveis do processo. Em particular, a análise de perfil é uma técnica estatística multivariada com medidas repetidas, geralmente utilizada em duas situações, onde compara-se as mesmas variáveis dependentes entre os grupos e quando existem várias medidas da mesma variável dependente. Assim sendo, essa análise pode ser feita por meio da análise de perfis através de um modelo univariado, de acordo com o planejamento do tipo split-plot, que impõe forte restrição quanto à matriz de variâncias-covariâncias, como por meio de um modelo multivariado, que utiliza uma matriz de variâncias-covariâncias sem restrições, chamada não-estruturada (SILVA, 2009).

De acordo com Adamczuc (2003), todos os grupos neste tipo de análise submetem-se sempre às mesmas condições experimentais, porém, é necessário testar três tipos de hipóteses, sendo elas: aceitar que os grupos apresentem perfis com uma mesma tendência de variação de comportamento; admitir que os perfis sejam paralelos e por último considerar que cada grupo em análise apresente igual resposta de comportamento nas diferentes condições experimentais. Portanto, estas hipóteses podem ser testadas em verdadeiras ou falsas por meio de um modelo estatístico específico chamado de análise de perfil. Porém, pelo fato deste modelo utilizar um número muito grande de operações algébricas matriciais e vetoriais, é quase impossível aplicar o mesmo, sem o auxílio de um software adequado. Neste caso, a análise multivariada de perfis foi realizada com o auxílio do software estatístico R versão 2.14.2.

Portanto, neste trabalho o nosso principal objetivo, é propor um procedimento adequado para a análise estatística de dados com medidas repetidas levando em consideração as restrições estatísticas inferenciais e a facilitação da aplicação das técnicas de análise.

A partir dos resultados obtidos, pode-se observar que o modelo de medidas repetidas no tempo ajustou-se adequadamente aos dados. Verificou-se para as hipóteses testadas $H_0^{(1)}$ (perfis horizontais), $H_0^{(2)}$ (perfis paralelos) e $H_0^{(3)}$ (perfis coincidentes) foram rejeita-

das ao nível de 5% de significância, ou seja, o crescimento da planta de girassol tende a ser melhor quando a inclusão de maiores níveis de matéria orgânica nas covas nas épocas de avaliação com o intervalo maior de dias. Compreendendo assim que a técnica abordada com medidas repetidas permite um efetivo acompanhamento do crescimento da planta do girassol.

2 Referencial teórico

2.1 Aspectos gerais sobre o girassol

O girassol é uma oleaginosa que está entre as cinco maiores culturas produtoras de óleo vegetal comestível do mundo, antecedido pelas culturas da soja, algodão, couve-nabiça e amendoim (OLIVEIRA et al. 2012). Este vegetal apresenta características agronômicas importantes, como maior resistência à seca, ao frio e ao calor, que a maioria das espécies normalmente cultivadas no Brasil. Apresenta ampla adaptabilidade a diferentes condições edafoclimáticas em razão de sua baixa sensibilidade fotoperiódica e seu rendimento é pouco influenciado pela latitude e pela altitude (GOMES et al. 2005).

A semente de girassol é uma das maiores fontes de nutrientes para a saúde mental e emocional. Isso porque essa semente oleaginosa contém um alto teor de vitamina B, fundamentais para o bem estar psicológico e vitamina B1, que ajuda a manter o cérebro alerta sob estresse. Além de possuir um óleo benéfico para o organismo, a semente de girassol é considerada uma importante fonte de fibras e proteínas com quantidades significativas de potássio, magnésio, cobre, fósforo, selênio e vitamina E, uma das mais importantes para o combate ao envelhecimento precoce. A demanda por sementes de girassol tem aumentado, devido à importância econômica do óleo extraído de seus aquênios, utilizado, principalmente, para consumo humano e animal, bem como matéria-prima para a produção de biocombustível (UNGARO, 2006).

Em Oliveira et al. (2012), ter conhecimento sobre a qualidade das sementes antes do plantio é o procedimento mais correto e seguro para se evitar aumentos no custo da lavoura. Pesquisas com o girassol, principalmente na área de controle de qualidade de sementes, são essenciais para o estabelecimento da cultura e se justificam pela eficiência da espécie. Logo, torna-se imprescindível ajustar tecnologias para a produção de sementes, que permitam homogeneizar o processo de germinação, bem como reduzir o tempo de germinação entre lotes, facilitando tomadas de decisão, em relação à origem e ao destino dos lotes.

De acordo com Oliveira (2013), três componentes vinculados e interativos, constituintes de um tipo de triângulo determinam o desempenho da semente: genética, ambiente e qualidade. Dos atributos que compõem a qualidade das sementes, o fisiológico se refere a sua capacidade de realizar sua função de propagação.

Segundo Gomes et al. (2005), a expansão da cultura do girassol pode ser prejudicada, entre outros fatores, pela qualidade fisiológica das sementes e pela presença de doenças causadas por vírus, bactérias e fungos. Sendo assim, diante do exposto, há necessidade de avaliar a qualidade física e fisiológica das sementes de girassol submetidas a diferentes tipos de tratamentos.

2.2 Fundamentos estatísticos

2.2.1 Medidas repetidas

Para Diggle (1988), Crowder e Hander (1990), o termo medidas repetidas é usada para nomear medidas dispostas da mesma variável ou na mesma unidade experimental em mais de uma ocasião. Dessa forma, diversas áreas de pesquisas utilizam planejamentos com medidas repetidas.

No entanto, estudos envolvendo planejamentos deste tipo, abrangem, entre outros, delineamentos com parcelas subdivididas (“*split-plot*”) em que casualizam-se os níveis de um fator (tratamentos) e depois casualiza os níveis do segundo fator em todos os níveis do primeiro; em delineamentos com intercâmbio (“*cross-over*”) as unidades experimentais recebem sequências de tratamento, ou seja, todos os tratamentos são aplicados às mesmas unidades experimentais, após agrupá-las em grupos de número igual ao número de tratamentos, além dos estudos longitudinais, que é um caso particular de medidas repetidas, quando há repetição da medida ao longo do tempo, (FERREIRA, 2012).

Em planejamentos longitudinais, a variável resposta pode ser contínua ou discreta, avaliada nas diferentes unidades experimentais, como, por exemplo, ganho de peso e número de folhas, que podem ser coligadas segundo tratamentos ou fatores. As unidades experimentais podem ser constituídas de um ou mais fatores ou tratamentos e cada uma dessas unidades pode gerar unidades observacionais, em que cada um desses conjuntos de observações pode ser entendido como um perfil individual de resposta para a variável em

estudo. O estudo desses perfis individuais pode ser um indício inicial para a inclusão de efeito aleatório no modelo (ROCHA, 2010).

Segundo Amador (2010), dados longitudinais são tidos como regulares quando o intervalo entre duas medidas consecutivas quaisquer for constante ao longo do tempo. E, se as observações forem feitas nos mesmos instantes de tempo em todas as unidades experimentais, tem-se uma estrutura balanceada em relação ao tempo.

Conforme Singer, Rocha e Nobre (2004), dados do tipo longitudinais, apresentam algumas vantagens, tais como:

- i) permitem estudar o comportamento da variável resposta média da unidade experimental sobre cada tratamento;
- ii) exige menos unidades experimentais (as medidas são feitas nas mesmas parcelas);
- iii) permite verificar existência de dependência da variável resposta em relação às covariáveis existentes;
- iv) diminuição do erro experimental (aumento do número de dados);
- v) melhora na precisão das estimativas dos parâmetros;

Rocha (2010), salienta ainda que, as observações sobre uma mesma unidade experimental tendem a ser correlacionadas, pela razão da obtenção sistemática dos dados longitudinais. E tal correlação pode ser modelada, explicada por meio da utilização de uma estrutura de covariâncias para os dados observados. E é imprescindível, que esse procedimento seja realizado, para que sejam válidas as inferências realizadas.

No estudo de medidas repetidas, analisada sob o enfoque da técnica de análise de perfil, Xavier e Dias (2001) alertam para o fato da validade dos resultados na análise de variância, da não violação dos pressupostos, de independência dos erros, pela casualização e homogeneidade, em que os dados estejam corretamente organizados (NEMEC, 1996 apud VIEIRA, 2006).

A análise de medidas repetidas no tempo pode ser verificada por meio da análise de perfis por intermédio de um modelo univariado, seguindo um planejamento do tipo “split-plot on time”, que impõe forte restrição quanto à matriz de variâncias-covariâncias, bem como por meio de um modelo multivariado que requer uma matriz sem restrições, ou não-estruturada, e o ajuste de curvas de crescimento pode ser realizado, por meio de

modelos mistos possibilitando o uso de diferentes tipos de estrutura para as matrizes de variâncias-covariâncias (FERREIRA 2012).

De modo geral, Vieira (2006) relata que umas das vantagens em utilizar o modelo multivariado (MANOVA) ao modelo univariado (ANOVA) é que além de ser uma alternativa para analisar dados mensurados no tempo, o mesmo não pressupõe uma estrutura específica para a matriz de covariância dos erros, apenas que os erros tenham distribuição normal multivariada. Contudo, os dois modelos exigem que a cada instante de tempo as variâncias e correlações sejam homogêneas.

Huynh e Feldt (1970) mostraram que, estudos com delineamento em esquema de parcelas subdivididas, pode ter a violação das pressuposições na análise de variância pelo fato da não aleatorização para os intervalos dos níveis do tempo, dessa forma, pode-se obter uma matriz de covariâncias com variâncias heterogêneas correspondentes aos erros das unidades experimentais, podendo resultar na probabilidade de rejeitar a hipótese nula, sendo a mesma verdadeira (erro tipo I).

De acordo com Vieira (2006), uma condição suficiente para que o teste F seja válido na análise de variância usual, em nível de subparcela, para o fator tempo e a interação \times tratamentos, é que a matriz de covariâncias seja igualmente correlacionada com variâncias iguais nas diferentes ocasiões, ou seja, tenha a forma de simetria composta.

A matriz de covariâncias Σ terá a forma de simetria composta, se e somente se, seus elementos da diagonal principal forem iguais a $(\sigma^2 + \sigma_1^2)$ e os demais elementos iguais a σ_1^2 , assim:

$$\Sigma = \begin{bmatrix} (\sigma^2 + \sigma_1^2) & \sigma_1^2 & \sigma_1^2 & \sigma_1^2 \\ \sigma_1^2 & (\sigma^2 + \sigma_1^2) & \sigma_1^2 & \sigma_1^2 \\ \sigma_1^2 & \sigma_1^2 & (\sigma^2 + \sigma_1^2) & \sigma_1^2 \\ \sigma_1^2 & \sigma_1^2 & \sigma_1^2 & (\sigma^2 + \sigma_1^2) \end{bmatrix}$$

em que:

σ^2 é a variância da sub parcela (intra-indivíduos) e,

σ_1^2 é a variância da parcela (entre indivíduos).

De modo geral, Vieira (2006) mostra que o requisito mais geral da forma de Σ é descrito por Huynh & Feldt (1970), denominada de Huynh-Feldt (H-F), no qual, determina que os elementos da matriz de covariâncias sejam expressos, para um $\lambda > 0$, da seguinte

forma;

$$\begin{aligned}\sigma_{ii} &= \sigma_i^2 \\ \sigma_{ij} &= [(\sigma_i^2 + \sigma_j^2) / 2] - \lambda\end{aligned}$$

em que:

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

É sabido que o requisito de H-F é imprescindível e suficiente, onde, por meio do teste F da análise de variância usual, no esquema de parcelas subdivididas no tempo, as variâncias da diferença entre os pares de erros sejam todas iguais. Dessa forma, o requisito é compatível à de simetria composta. Entretanto, caso a matriz de covariância não atenda ao requisito de H-F, mas possua uma forma muito próxima, que será aferida pelo teste de Mauchly, ainda assim as condições do teste F, não se alteram.

De acordo com Ferreira (2012), nas situações em que as pressuposições de normalidade, de independência e do requisito H-F para a matriz de covariâncias não serem satisfeitas, à análise de variância multivariada ou análise multivariada de perfil é uma alternativa para realizar as análises, porque exigem apenas que os erros tenham distribuição normal multivariada e não pressupõe uma estrutura específica de matriz de covariância dos erros.

Em determinadas ocasiões a matriz de covariância dos erros não satisfaz aos requisitos de H-F, porém, pode ser que haja uma aproximação estatística. Dessa forma, Mauchly (1940) sugeriu um teste de esfericidade com a finalidade de verificar se uma determinada população normal multivariada possui variâncias iguais e correlações nulas, se o teste for positivo a população será denominada de esférica, de outro modo, aconselha-se o uso da análise multivariada de perfis.

2.2.1.1 Teste de esfericidade de Mauchly

Proposto por Mauchly (1940), o teste de esfericidade consiste em verificar se uma população normal multivariada possui variâncias iguais e correlações nulas. Se esse requisito for atendido, a matriz de covariâncias será denominada de esférica. Dessa forma, sua aplicação compreende a utilização da matriz de contrastes ortonormais, transformando a matriz de covariância original para uma forma ortonormalizada, cuja esfericidade será avaliada pelo teste e, sendo esférica a matriz original, o requisito de H-F será satisfeito, comprometendo a validade do teste F. Por meio do requisito de H-F para a matriz de covariâncias das medidas repetidas dos elementos atribuídos nos $(t-1)$ contrastes ortogonais normalizados. A ortogonalidade dos contrastes certifica que:

- i) cada contraste está agregado a uma única parcela da variabilidade explicada pelo efeito que se está testando;
- ii) testa-se o número máximo de hipóteses, em que cada hipótese é agregada a uma única parcela da variabilidade explicada pelo modelo;
- iii) o teste é aproximadamente independente.

De acordo com Vieira (2006), se Σ é a matriz de covariâncias das medidas repetidas no tempo e houver t tempos, mais de um conjunto de $(t-1)$ contrastes ortogonais, pelo menos um contraste ortogonal será normalizado quando for dividido por sua norma euclidiana. Logo, o teste de hipótese para o requisito atribuído pelo teste H-F para as covariâncias dos contrastes é dado pela seguinte expressão:

$$H_0 : \mathbf{C}\Sigma\mathbf{C}' = \lambda\mathbf{I}$$

em que,

Representando o total de hipóteses nulas, \mathbf{C} é a matriz $(t-1) \times t$ dos coeficientes dos contrastes ortogonais normalizados; Σ é a matriz de covariâncias $(t \times t)$; λ é um escalar maior do que zero e \mathbf{I} é a matriz identidade $(t-1) \times (t-1)$. Assim, cumprindo este requisito, a matriz de covariâncias Σ é denominada de esférica.

Ferreira, (2012) salienta que, a estatística W do teste Mauchly (1940) para validade da hipótese referente ao requisito de esfericidade é dado pela seguinte expressão:

$$W = \frac{(t-1)^{t-1} |\mathbf{C}\Sigma\mathbf{C}'|}{(tr(\mathbf{C}\Sigma\mathbf{C}'))^{t-1}}$$

em que:

\mathbf{S} é a matriz de covariâncias amostrais, de dimensões $(t \times t)$ e tr é o traço. Sendo, nas t medidas repetidas escolhidos $(t-1)$ contrastes ortogonais normalizados, e sabendo que as linhas da matriz $\mathbf{C}_{(t-1) \times t}$ seguem esse mesmo padrão, calcula-se, então a matriz $\mathbf{C}\Sigma\mathbf{C}'_{(t-1) \times (t-1)}$ (KIRK, 1995 apud VIEIRA, 2006).

Todavia, o valor de \mathbf{W} não depende da escolha particular para esta matriz de transformação \mathbf{C} . Se Σ é estimada com $v = (t-1)$ graus de liberdade e se considerar a definição de um escalar para uma melhor exatidão da aproximação pelo Qui-quadrado, temos:

$$\gamma = v - \frac{2t^2 - 3t + 3}{6(t-1)}.$$

Dessa forma, a estatística W para a hipótese nula $H_0 : \mathbf{C}\Sigma\mathbf{C}' = \lambda\mathbf{I}$, poderá ser averiguada pela expressão

$$\chi^2 = -\lambda \ln(W),$$

Considere $f = \frac{1}{2}t(t-1) - 1$ graus de liberdade, então se, $-\lambda \ln(W) > \chi^2_{(f;\alpha)}$, há fortes evidências de não aceitar a hipótese nula. Embora, Freitas et al. (2008) não recomendem o uso deste teste quando há dúvidas da pressuposição de normalidade dos erros, devido à sua sensibilidade a ausência de normalidade e, com isso, acarreta a diminuição de seu poder. Sendo assim, é aconselhável a utilização de testes de normalidade para analisar essa pressuposição.

Segundo Malheiros (2004), quando utilizar a análise univariada em esquema de parcelas subdivididas no tempo e a matriz de covariância dos erros intraindivíduos não possuir, ao menos, a forma de H-F, uma alternativa é proceder ao ajuste dos graus de liberdade do teste F para os fatores presentes na subparcela, porque até então, existem fortes razões que favorecem a suposição de esfericidade. De outro modo, um teste F modificado, ou um teste multivariado ou até mesmo, uma abordagem não paramétrica deve ser aplicada. Ainda assim, as decisões sempre devem ser tomadas com base nos dados.

2.2.1.2 Análise multivariada de perfis

O modelo para análise multivariada de perfis são equivalentes aos costumeiramente considerados em análise de variância (MANOVA), realizada por meio dos planejamentos do tipo parcela subdividida no tempo, além disso, tem a vantagem de favorecer uma melhor compreensão. Nesses casos, o tempo é tomado como uma variável discreta, ordenando a variável resposta na parcela experimental. De acordo com Xavier (2000), esse modelo pode ser descrito matricialmente da seguinte forma:

$$\mathbf{Y} = \mathbf{X}\beta + \epsilon$$

em que,

$\mathbf{Y}_{(N \times t)} = [y_{11}, \dots, y_{gn_g}]'$ é a matriz dos dados observados $N \times t$ de t respostas para os N indivíduos;

$y_{ij} = [y_{ij1}, \dots, y_{ijt}]'$ é o perfil de respostas da unidade experimental (i, j) ;

$$\mathbf{X}_{(N \times g)} = \begin{bmatrix} 1_{n1} & 0 & \dots & 0 \\ 0 & 1_{n2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1_{ng} \end{bmatrix}$$
 é a matriz de delineamento conhecida, ou seja, de especificação do modelo;

$$\beta_{(g \times t)} = \begin{bmatrix} \mu'_1 \\ \mu'_2 \\ \vdots \\ \mu'_g \end{bmatrix}$$
 é a matriz de parâmetros;

$\varepsilon_{(N \times t)} = [\varepsilon_{11}, \dots, \varepsilon_{gn_g}]'$ é a matriz do erro experimental.

Dessa forma, esse modelo assume que ε seja uma matriz de variáveis aleatórias independentes, onde as linhas são não correlacionadas e que segue uma distribuição normal multivariada com a média $\mathbf{0}$ e matriz de covariâncias Σ , em que, Σ é chamada não estruturada e tem $\frac{t(t+1)}{2}$ parâmetros. Com base nisto, supõe-se que os N perfis de respostas $y_{ij} \sim N_t(\mathbf{X}\beta, \Sigma)$, em que as matrizes de covariâncias correspondentes sejam todas iguais e sigam a forma geral:

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1t} \\ \sigma_{12} & \sigma_2^2 & \dots & \sigma_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1t} & \sigma_{2t} & \dots & \sigma_t^2 \end{bmatrix}$$

Armas (2004), salienta que um dos principais benefícios obtidos pelo uso da MANOVA, se dá pelo fato da matriz de variância e covariância não necessariamente deva obter uma forma estruturada, descartando o aspecto de esfericidade e, assim sendo, toda a importância relacionada ao teste F e correções dos graus de liberdade. Este acontecimento se dá pelo fato da MANOVA exigir um erro específico para contrastes com 1 grau de liberdade, nos quais, cada contraste está vinculado com seu termo de erro específico.

Lima (s.d), acentua algumas restrições existentes na aplicação da análise multivariada de perfis, tais como:

1. a análise só pode ser realizada quando se têm um número de unidades experimentais maior que o número de situações ($N > t$);

2. ausência de perfis de dados completos (na perda qualquer das observações para um mesmo indivíduo, seu perfil de resposta será excluído da análise);
3. o poder dos testes é pequeno;
4. as distintas estatísticas dos testes pode resultar em conclusões desiguais.

Por meio do uso da análise multivariada de perfis, têm-se como objetivo o interesse em testar hipóteses referente aos valores médios da variável resposta, nas diversas situações de observação (tempo), com a finalidade de responder as questões descritas a seguir nas hipóteses de interesse.

- i) Hipótese de perfis paralelos - os perfis médios de resposta são paralelos? Se sim, a interação entre tratamentos e o efeito do tempo é nula.
- ii) Hipótese de perfis coincidentes - os perfis médios de resposta correspondentes aos diversos tratamentos são coincidentes? Se sim, sendo os perfis paralelos, o efeito de tratamento é nulo.
- iii) Hipótese de perfis horizontais - os perfis médios de resposta relacionado aos diferentes tratamentos são paralelos ao eixo das abcissa? Se sim, sendo os perfis paralelos, o efeito do tempo é nulo.

Logo, as hipóteses sobre os parâmetros a serem testadas também podem ser expressas na forma da hipótese linear geral:

$$H : \mathbf{CBT} = 0$$

onde $C_{(c-1) \times (c+b+1)}$ e $T_{t \times (t-1)}$ são matrizes de constantes conhecidas com postos c e t , respectivamente. Tem-se que a matriz \mathbf{C} é responsável por comparações entre os grupos (linhas da matriz \mathbf{B}), e a matriz \mathbf{T} é responsável por comparações entre os tempos (colunas da matriz \mathbf{B}). Ainda pode ter-se possíveis correspondências obtidas através da hipótese linear geral, onde, Andrade e Singer 1986, ressaltam que os testes da hipótese geral são provenientes de inúmeros critérios. No geral, as estatísticas dos testes são funções das raízes características da matriz \mathbf{HE}^{-1} , onde

$$\begin{aligned} \mathbf{H} &= \mathbf{T}'((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y})\mathbf{C}'[\mathbf{C}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{C}']^{-1}\mathbf{C}((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y})\mathbf{T} \\ \mathbf{E} &= \mathbf{T}'\mathbf{Y}'[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{Y}\mathbf{T} \end{aligned}$$

em que,

\mathbf{H} é a matriz de somas de quadrados e produtos cruzados devido a hipótese nula.

\mathbf{E} é a matriz de somas de quadrados e produtos cruzados devido ao erro.

De acordo com Xavier (2000), diversas estatísticas de testes para testar as hipóteses de ausência de efeito de tratamento, interação e tempo, para o caso multivariado, são obtidas por meio da união-intersecção de Roy e da razão de verossimilhança de Wilks, estas estatísticas são dados por:

Lambda de Wilks:

$$\Lambda = \prod_{i=1}^p (1 + \lambda_i)^{-1}$$

Traço de Pillai:

$$v = \sum_{i=1}^p \frac{\lambda_i}{(1 + \lambda_i)}$$

Traço de Lawley-Hotelling:

$$U = \sum_{i=1}^p \lambda_i$$

Maior raiz característica de Roy:

$$\theta = \frac{\lambda_i}{1 + \lambda_i},$$

em que λ_i são os autovalores obtidos da solução da raiz característica de \mathbf{HE}^{-1} .

Em Vieira (2006), a aplicação desses testes não exigem que as matrizes de covariância satisfaçam a condição de H-F, embora os casos multivariados tem por base uma matriz não estruturada. Logo, as estimativas da matriz de covariância podem ser obtidas pelo método dos momentos. Contudo, os quatro testes multivariados podem resultar em diversos níveis descritivos. Em geral, o teste mais usado é o Lambda de Wilks, por apresentar o maior poder do teste.

Segundo Xavier (2000), a desvantagem no uso da análise multivariada, refere-se a insuficiência do poder para estimar os parâmetros da matriz de covariâncias, isto ocorre quando t (número de ocasiões medidas ou tempos) é grande e N é pequeno.

As estatísticas dos quatro teste MANOVA têm a mesma probabilidade de rejeitar H_0 quando a mesma é verdadeira, caso contrário, quando H_0 é falsa, os testes têm diferentes probabilidades de rejeição. Sendo assim, o teste de Roy é mais poderoso que os outros, se os vetores médios forem colineares, já que ele usa apenas o maior autovalor,

mas, do contrário os outros três testes têm mais poder que o de Roy. Porém, o teste Lambda de Wilks permanece dominante, por causa de sua exigibilidade. (RENCHE E CHRISTENSEN, 2012).

3 Material e Métodos

Nesta seção encontram-se as principais metodologias que servirão de base para este estudo, no que diz respeito a utilização do modelo multivariado com medidas repetidas e os dados utilizados neste trabalho.

3.1 Caracterização dos dados

Com a intenção de aplicar a teoria abordada neste trabalho, foi utilizado um experimento em blocos completos casualizados que contém um fator com 6 níveis, 4 repetições e avaliados em 5 épocas, resultando numa amostra de 120 observações. Para o experimento adotado com 4 blocos, considera-se o diâmetro do caule (mm) a variável resposta, quantidade de matéria orgânica (esterco de caprinos) o tratamento e ($C_0 = 0, C_1 = 1, C_2 = 2, C_3 = 3, C_4 = 4, C_5 = 5$) os seis níveis em litros por cova. Como critério de estudo foi utilizado as medidas do diâmetro do caule do girassol ao longo do crescimento da planta durante 5 épocas de avaliação aplicadas ao experimento em 30, 45, 60, 70 e 80 dias após a emergência, relacionado com os 6 níveis de quantidade de matéria orgânica. Para análise destes dados foi empregado o *Software* livre R versão 2.14.2.

Dessa forma, o delineamento experimental no qual as respostas foram coletadas para cada unidade experimental, onde acredita-se que as medidas ao longo do tempo devem estar correlacionadas, caracteriza uma análise multivariada com delineamento de medidas repetidas. A Tabela 1 apresenta a relação das 120 observações com suas respectivas medidas durante as 5 épocas de avaliação.

Tabela 1: Relação das medidas do diâmetro do caule do girassol (mm) em cada época de avaliação

Tratamento	Bloco	Avaliação 1	Avaliação 2	Avaliação 3	Avaliação 4	Avaliação 5
C_0	1	7,09	8,97	12,52	14,31	14,16
C_0	2	6,93	9,06	15,31	13,72	13,48
C_0	3	8,04	10,23	12,49	13,67	14,60
C_0	4	8,20	8,56	12,04	14,27	13,91
C_1	1	9,45	11,62	18,82	18,82	19,10
C_1	2	9,34	12,60	16,65	19,10	18,62
C_1	3	8,90	11,88	17,92	19,36	19,01
C_1	4	10,45	14,36	16,92	17,08	18,56
C_2	1	10,82	13,20	21,35	23,45	23,50
C_2	2	11,20	13,64	20,33	23,64	23,40
C_2	3	11,20	14,23	19,94	23,45	24,10
C_2	4	11,05	14,08	20,66	24,10	23,10
C_3	1	12,32	14,44	25,30	26,33	27,45
C_3	2	12,12	15,90	25,54	27,02	25,12
C_3	3	13,80	15,70	20,89	25,92	26,30
C_3	4	12,85	17,30	23,20	26,40	26,20
C_4	1	13,86	17,02	25,38	28,06	27,08
C_4	2	14,58	18,22	23,84	27,08	26,82
C_4	3	14,33	17,20	24,93	27,55	27,58
C_4	4	13,94	18,10	25,87	28,10	27,44
C_5	1	13,66	17,33	23,20	28,05	28,30
C_5	2	15,20	17,98	24,12	27,23	28,08
C_5	3	14,89	18,05	24,85	27,50	26,33
C_5	4	15,03	18,20	25,36	26,91	26,60

3.2 Métodos utilizados

A princípio, é notório que as análises estatísticas para um conjunto de dados com medidas repetidas devem satisfazer as seguintes condições: normalidade, independência, homogeneidade de variâncias entre grupos e circularidade da matriz de covariância conjugada (TOUTENBURG, 2002). Dessa forma, deve-se adotar o procedimento de análise na seguinte ordem:

- i) Verificar as suposições de normalidade e independência do conjunto de dados:

Normalidade multivariada por χ^2 e gráfico gama (γ).

- ii) Aplicar um teste para verificar a homogeneidade das variâncias:

$$H_0 = \Sigma_1 = \Sigma_2 = \Sigma_3 \dots \Sigma_g.$$

iii) Testar a condição de circularidade da matriz de covariância conjugada:

$$H_0 = \Sigma = \sigma^2[(1 - \rho)\mathbf{I} + \rho\mathbf{J}] \quad \text{ou} \quad H_0 = \mathbf{C}\Sigma\mathbf{C}' = \mathbf{I}\sigma^2.$$

iv) Testar as hipóteses sobre os parâmetros do modelo via MANOVA ou ANOVA.

De acordo com Toutenburg (2002), uma das exigências da análise multivariada (MANOVA) é que a homogeneidade das matrizes de covariâncias sejam satisfeitas. Desse modo, para testar tal requisito usa-se o teste de Bartlett generalizado por Box (1949) para comprovar a adequação do modelo para igualdade de variâncias, ou seja, $H_0 = \Sigma_1 = \Sigma_2 = \Sigma_3 \dots \Sigma_g$, embora a estatística M de Box é αM , dada pela expressão:

$$M = (n - g) \ln |S| - \sum_{k=1}^g v_k \ln |S_k|,$$

$$\alpha = 1 - \left[\frac{2p^2 + 3p - 1}{6(p+1)(g-1)} \left(\sum_{k=1}^g \frac{1}{v_k} - \frac{1}{n-g} \right) \right]$$

em que, $v_k = n_k - 1$ os graus de liberdade associados a cada grupo com amostra de tamanho n , S_k matriz de variâncias e covariâncias do grupo k , $k = 1, 2, \dots, g$, e \mathbf{S} é a matriz de covariância amostral combinada.

Assim, sob a hipótese H_0 , temos a seguinte distribuição aproximada

$$\alpha M \sim \chi^2_{\left[\frac{1}{2}p(p+1)(g-1)\right]}.$$

Portanto, se a hipótese H_0 não for rejeitada, é porque existem evidências de que as matrizes de covariâncias são homogêneas e a MANOVA pode ser realizada.

Dessa forma, verifica se a matriz de covariâncias atende à condição H-F esfericidade (não-significativa) pelo teste de esfericidade de Mauchly, concluindo que a matriz é esférica e o experimento pode ser analisado univariadamente como se fosse um experimento em parcela subdividida. Caso o teste resulte em significativo com $p - \text{valor} < 0,05$ para as unidades observadas, o procedimento pode ser realizado com a análise multivariada de medidas repetidas, mais conhecida por MANOVA.

Entretanto, sendo as matrizes de covariâncias homogêneas, pode-se conduzir com o teste de esfericidade proposto por Mauchly (1940) para a hipótese $H_0 : \mathbf{C}\Sigma\mathbf{C}' = \lambda\mathbf{I}$, de acordo com a seguinte expressão:

$$X^2 = -\gamma \ln(W),$$

a violação da condição de esfericidade refere-se sobre a rejeição da hipótese H_0 , onde, se $-\gamma \ln(W) > \chi_{(f;\alpha)}^2$, rejeita-se a hipótese nula e podemos afirmar que existem fortes evidências para a validade da análise multivariada.

3.2.1 Hipóteses testadas

Aplicando-se a análise multivariada para esse caso, as hipóteses de interesse a serem testadas são as seguintes:

- $H_0^{(1)}$: Os perfis médios de resposta correspondentes aos diversos tratamentos são horizontais, isto é, não existe efeito do fator EA (Época de Avaliação) tempos, em relação aos parâmetros do modelo:

$$H_0^{(1)} : \begin{bmatrix} \mu_{11} \\ \mu_{21} \\ \vdots \\ \mu_{n1} \end{bmatrix} = \dots = \begin{bmatrix} \mu_{1t} \\ \mu_{2t} \\ \vdots \\ \mu_{nt} \end{bmatrix}$$

em que, rejeita-se a hipótese nula, quando existir efeito no fator EA (Época de Avaliação).

- $H_0^{(2)}$: os perfis médios de resposta correspondentes são paralelos, isto é, não existe interação entre o fator EA \times tratamento. Na forma matricial, em relação aos parâmetros do modelo:

$$H_0^{(2)} : \begin{bmatrix} \mu_{11} - \mu_{12} \\ \mu_{12} - \mu_{13} \\ \vdots \\ \mu_{1(t-1)} - \mu_{1t} \end{bmatrix} = \dots = \begin{bmatrix} \mu_{g1} - \mu_{g2} \\ \mu_{g2} - \mu_{g3} \\ \vdots \\ \mu_{n(t-1)} - \mu_{nt} \end{bmatrix}$$

em que, rejeita-se a hipótese nula quando houver interação entre o fator EA e o fator tratamento.

- $H_0^{(3)}$: os perfis médios de resposta correspondentes aos diversos tratamentos são coincidentes, isto é, não existe diferença entre o fator tratamento. Na forma matri-

cial, em relação aos parâmetros do modelo:

$$H_0^{(3)} : \begin{bmatrix} \mu_{11} \\ \mu_{12} \\ \vdots \\ \mu_{1t} \end{bmatrix} = \dots = \begin{bmatrix} \mu_{g1} \\ \mu_{g2} \\ \vdots \\ \mu_{gt} \end{bmatrix}$$

. em que, rejeita-se a hipótese nula quando houver diferença do fator tratamento.

As estatísticas para testar as hipóteses de ausência de efeito de tratamento, interação e tempo podem ser conferidas por meio dos testes Lambda de Wilks, Traço de Pillai, Traço de Hotelling-Lawley, como também o Teste de Roy, já relatadas na seção anterior.

3.2.2 Procedimento da análise

As análises sugeridas para o experimento em blocos completos casualizados, foram realizadas por meio da utilização do *Software R* versão 2.14.2, em que possui uma estrutura de fácil compreensão para obtenção dos resultados, tanto para o modelo univariado quanto para o modelo multivariado.

4 Resultados e Discussão

Nesta seção encontra-se os resultados, aos quais seguem a formulação sugerida nos objetivos do trabalho. À princípio a idéia de medidas repetidas faz sentido ao tema deste estudo porque as medidas do diâmetro do caule do girassol (mm) está relacionada a época de avaliação ao longo do tempo.

Assim como Armas (2004), Ferreira (2012) e Barbosa (2009), também foi utilizado o Software R para verificar as análises, usando os pacotes necessários *lattice*, *psych*, *car* e *nlme* para fazer o teste de esfericidade de Mauchly e os testes multivariados produzidos por: *Lambda de Wilks*, *Traço de Pillai*, *Traço de Hotelling-Lawley* e *Roy*, sendo a interação entre os fatores entre e intra-indivíduos automaticamente incluídas no modelo, tanto para análise univariada ou multivariada.

Primeiramente para que uma análise estatística tenha validade e confiabilidade ao utilizar um modelo com medidas repetidas, deve-se verificar as pressuposições do modelo e sua adequação ao modelo. Sendo assim, foi realizado o teste de normalidade de Shapiro-Wilk para cada época de avaliação em estudo, bem como o teste de homogeneidade das variâncias pelo teste de Bartlett. Entretanto, observa-se que as variações ocorreram dentro do esperado, mantendo-se a precisão das análises e ainda busca-se verificar os valores médios do desempenho do tratamento em cada época de avaliação.

Afim de verificar a normalidade dos dados, aplicou-se o teste de Shapiro-Wilk, no qual foi bastante significativo, logo a hipótese de nulidade não foi rejeitada ao nível de 5% de significância, considerando assim a normalidade dos dados. Também verificou-se a homogeneidade de variâncias dentro de cada avaliação por meio do teste de Bartlett com o resultado de 64,14 com 9 graus de liberdade e valor p menor que 0, em que foi bastante significativo, assim a hipótese de nulidade não foi rejeitada para praticamente todas as observações, porém ressalva-se que para uma única observação há rejeição da hipótese nula.

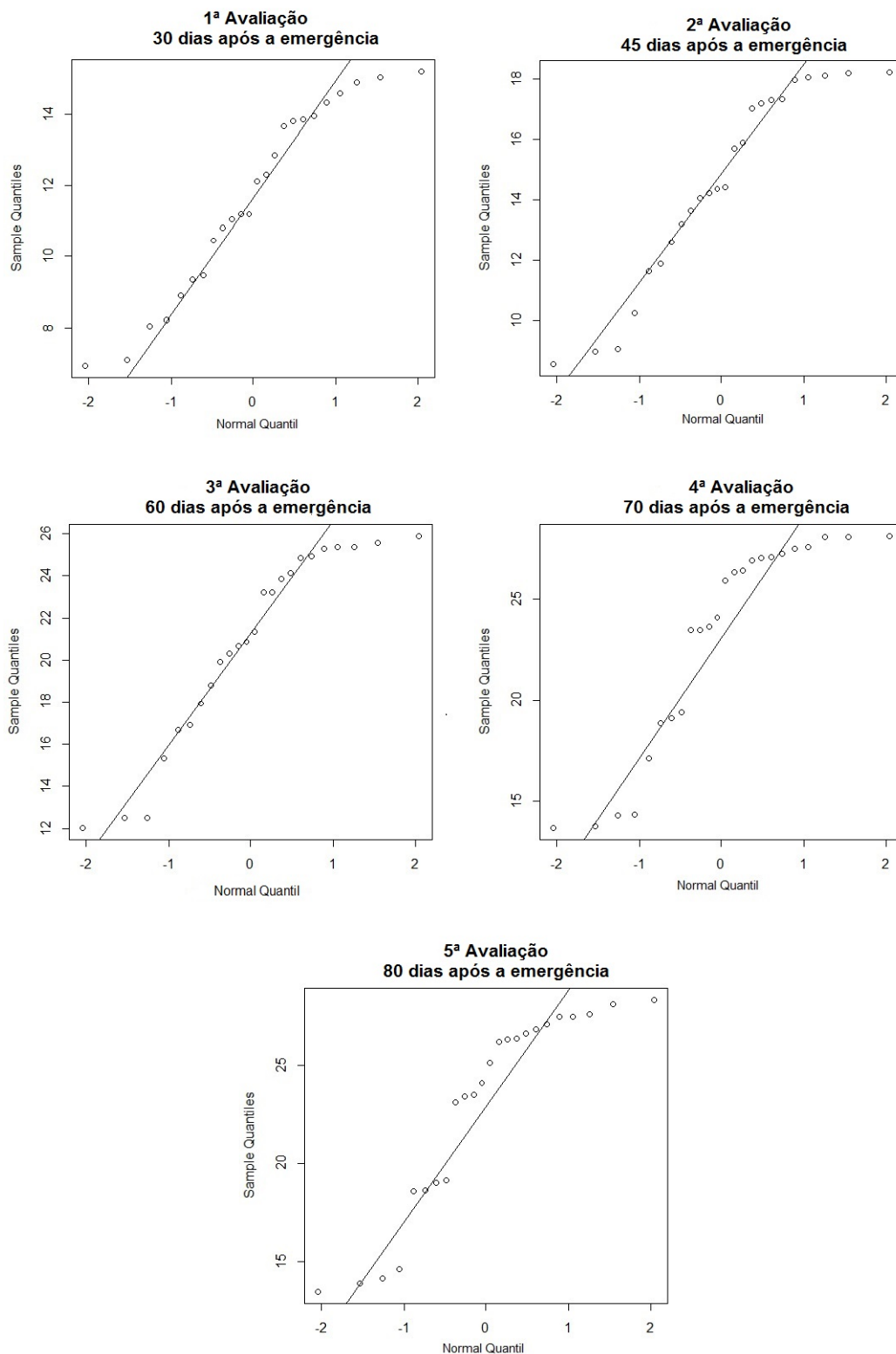


Figura 1: Gráfico qq-plot de cada época de avaliação ao longo do ciclo do crescimento das plantas

O teste de Shapiro-Wilk foi bastante significativo exceto para a época de avaliação 1, onde pode-se comprovar por meio dos gráficos de resíduos (ver figura 1), já que a hipótese

de nulidade não foi rejeitada ao nível de 5% de significância, confirmando a normalidade dos dados para 4 épocas de avaliação. O teste de homogeneidade de variâncias dentro de cada época de avaliação também foi bastante significativo, em que a hipótese de nulidade não foi rejeitada para praticamente todas as observações, porém ressalva-se que para uma única observação a rejeição da hipótese nula.

Tabela 2: Teste Shapiro-Wilk (W) para cada época de avaliação

Época de avaliação	W	Valor p
Avaliação 1	0,93	0,121
Avaliação 2	0,91	0,032
Avaliação 3	0,89	0,016
Avaliação 4	0,83	0,0009
Avaliação 5	0,85	0,002

* significativo ao nível de 5% de probabilidade

Nas Tabelas 3 e 4 de variância-covariância e de correlação entre as épocas de avaliação, onde verifica-se que, as variâncias não são homogêneas entre cada época de avaliação e existe correlação entre as medidas do diâmetro do caule do girassol com a época de avaliação. Sendo assim, as matrizes de variância-covariância e correlação não atendem a condição H-F (simetria composta). Logo, comprova-se que a análise estatística deverá ser multivariada com medidas repetidas.

Tabela 3: Variância-Covariância de cada época de avaliação

	Avaliação 1	Avaliação 2	Avaliação 3	Avaliação 4	Avaliação 5
Avaliação 1	700,44	816,19	1069,17	1286,396	1259,74
Avaliação 2	816,19	1011,14	1297,73	1536,62	1497,08
Avaliação 3	1069,17	1297,73	2021,85	2254,16	2153,91
Avaliação 4	1286,39	1536,62	2254,16	2706,48	2596,28
Avaliação 5	1259,74	1497,08	2153,01	2596,25	2544,10

Tabela 4: Correlação de cada época avaliação

	Avaliação 1	Avaliação 2	Avaliação 3	Avaliação 4	Avaliação 5
Avaliação 1	1	0,97	0,89	0,93	0,94
Avaliação 2	0,97	1	0,91	0,93	0,93
Avaliação 3	0,89	0,91	1	0,96	0,95
Avaliação 4	0,93	0,93	0,96	1	0,99
Avaliação 5	0,94	0,93	0,95	0,99	1

Na Tabela 5 e 6 encontra-se os quatro testes multivariados *Lambda de Wilks*, *Traço de Pillai*, *Traço de Hotelling-Lawley* e *Roy* para época de avaliação e para a interação época de avaliação \times níveis de matéria orgânica, sendo todos independentes da estrutura da matriz de covariância.

Pelos testes multivariados, testou-se a hipótese nula de perfis horizontais, ou seja, que testa o efeito da época de avaliação e de perfis paralelos, ou seja, para testar a interação entre a época de avaliação \times níveis de matéria orgânica. Para a fonte de variação que foi testada, foi obtida uma matriz de soma de quadrados e produtos, além da matriz de soma de quadrados e produtos referentes aos erros do modelo. As estatísticas de teste foram obtidas a partir das raízes latentes da matriz. Contudo, foi usado a matriz da soma de quadrados e produtos para a época de avaliação e a matriz soma de quadrados e produtos dos erros. Além disso, foram impressos os valores, e, utilizados para obter a aproximação com a estatística F.

Tabela 5: Resultado da análise multivariada, testando o efeito da época de avaliação.

Estatística	Valor	F	G.L. Num ¹	G.L. Den ²	Pr > F
Lambda de Wilks	0,004	919,21	4	15	< 0,001*
Traço de Pillai	0,99	919,21	4	15	< 0,001*
Traço de Hotelling-Lawley	245,12	919,21	4	15	< 0,001*
Raiz máxima de Roy	245,12	919,21	4	15	< 0,001*

* significativo ao nível de 5% de probabilidade

Tabela 6: Resultado da análise multivariada, testando o efeito da interação da época de avaliação \times níveis de matéria orgânica.

Estatística	Valor	F	G.L. Num ¹	G.L. Den ²	Pr > F
Lambda de Wilks	0,005	3,75	20	50,69	< 0,001*
Traço de Pillai	1,27	1,68	20	72	0,057
Traço de Hotelling-Lawley	13,24	8,94	20	54	< 0,001*
Raiz máxima de Roy	12,8	46,1	20	18	< 0,001*

* significativo ao nível de 5% de probabilidade

Conforme as tabelas 5 e 6 podemos verificar a rejeição da hipótese $H_0^{(1)}$ para o efeito da época de avaliação (perfis horizontais), o que indica que existe dependência entre o tempo apropriado de tratamento para o crescimento do diâmetro do caule do girassol. Contudo, para a interação época de avaliação \times níveis de matéria orgânica (perfis paralelos), há evidências para a rejeição da hipótese $H_0^{(2)}$, sendo assim, podemos afirmar que esses fatores são dependentes um do outro, ou seja, existe uma relação entre a época de avaliação com o nível de matéria orgânica para o crescimento do diâmetro do caule do girassol. Pode-se concluir que pela análise multivariada, não houve ocorrência de perfis horizontais, paralelos e coincidentes.

Para identificar os intervalos de tempo entre cada época de avaliação, foi construído tabelas ANOVAs para cada contraste.

Com base na tabela 7, pode-se verificar que os intervalos de contrastes Avaliação 2 *vs*

Tabela 7: Soma de quadrados (S.q.), graus de liberdade (g.l.), Média quadrática (M.q.), valor F e *p*-valor de contrastes formados entre as épocas de avaliação.

Contrastes	g.l.	S.q.	M.q.	F	<i>Pr</i> > <i>F</i>
Avaliação 2 <i>vs</i> avaliação 1	5	796,5	159,3	2,79	< 0,001*
Avaliação 3 <i>vs</i> avaliação 2	5	4680	936	3,13	< 0,001*
Avaliação 4 <i>vs</i> avaliação 3	5	2011,1	402,2	2,37	0,08
Avaliação 5 <i>vs</i> avaliação 4	5	108,97	21,79	0,32	0,89

* significativo ao nível de 5% de probabilidade

avaliação 1 e Avaliação 3 *vs* avaliação 2 tiveram efeito significativo, esta análise apontou que os intervalos de contrastes Avaliação 4 *vs* avaliação 3 e Avaliação 5 *vs* avaliação 4 não foram significativos, ou seja, não há efeito em relação a época de avaliação. Para manter um $\alpha = 0,05$, um ajuste de Bonferroni deve ser efetuado de acordo com o número de testes avaliados. Para o nosso caso um $\alpha = 0,05/4 = 0,0125$ para cada contraste poderia ser utilizado.

Por meio das figuras 2, 3 e 4 compreendemos melhor essas relações que correspondem aos perfis individuais do diâmetro médio em relação a época de avaliação em cada tratamento aplicado as plantas de girassol.

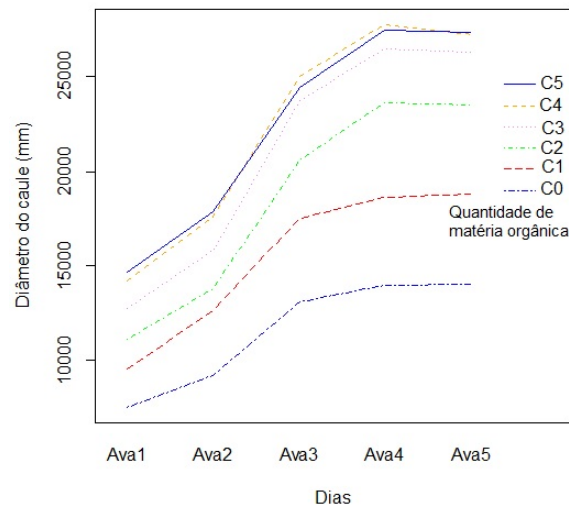


Figura 2: Perfil do diâmetro médio do caule do girassol (mm) em relação às épocas de avaliação em cada tratamento

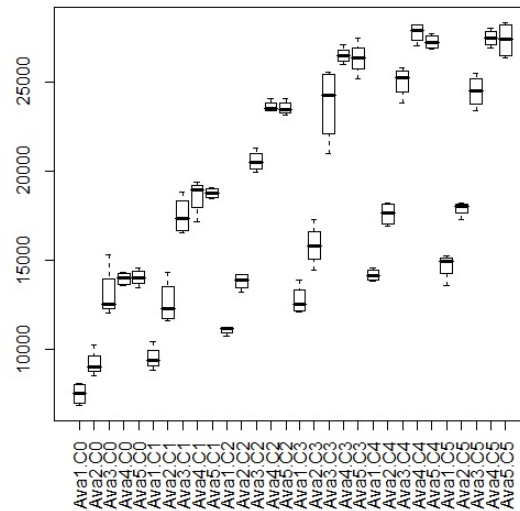


Figura 3: Boxplot dos dados originais

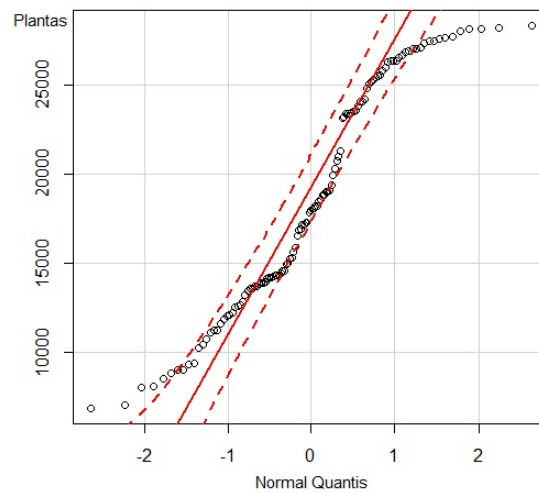


Figura 4: QQ-plot dos dados originais

Nota-se, que nesse experimento o aumento das medidas do diâmetro do caule do girassol ocorre na medida que o nível de matéria orgânica (esterco de caprinos) aplicado nas covas for em maior quantidade, juntamente com o intervalo maior de dias da época de avaliação.

5 Conclusões

De forma objetiva, esse trabalho teve resultados satisfatórios a contribuir positivamente, na tentativa de expor a análise multivariada de perfil como uma ferramenta de extrema importância quando se têm dados com medidas repetidas, aos quais, permite confirmar a eficácia das medidas de melhoria adotadas.

Dessa forma, por meio dos resultados alcançados nesse estudo, pode-se concluir que o modelo de medidas repetidas no tempo ajustou-se adequadamente aos dados, em que as medidas tomadas ao longo do tempo na mesma unidade experimental se correlacionam entre si, justificando uma análise multivariada. As pressuposições do modelo foram validadas, o que tornou a análise estatisticamente significativa. Para as hipóteses testadas $H_0^{(1)}$ (perfis horizontais), $H_0^{(2)}$ (perfis paralelos) e $H_0^{(3)}$ (perfis coincidentes) foram rejeitadas ao nível de 5% de significância, concluindo assim, que cada época de avaliação e o nível de matéria orgânica têm influência quanto ao tamanho das medidas do diâmetro do caule do girassol aqui estudadas. Portanto, a técnica abordada com medidas repetidas permite um efetivo acompanhamento do crescimento da planta do girassol.

Logo, dada a significância estatística da interação tratamento x época de avaliação pode-se concluir que o comportamento médio do diâmetro do caule de girassol é diferente na presença de cada quantidade de esterco usada na adubação.

Referências

ADAMCZUK, P. C. **Análise de perfis: Aplicação a um caso real na indústria metalúrgica.** 218 p. Dissertação (Métodos Numéricos em Engenharia-Programação Matemática) - Universidade Federal do Paraná, Curitiba, julho, 2003.

ANDRADE, D. F.; SINGER, J. M. **Análise de dados longitudinais.** In: VII Simpósio nacional de Probabilidade e Estatística - SINAPE. Campinas, São Paulo: [s.n.], 1986. p. 106.

AMADOR, J. P. **Modelos mistos no ajuste de curvas de crescimento de *Ricinus communis* L.** 2010. 139 p. Tese (Doutorado em Agronomia) - Universidade Federal de Santa Maria, Santa Maria, 2010.

ARMAS, E. D. **Análise de medidas repetidas.** 2004.

BENINCASA, M. M. P. **Análise de crescimento de plantas (noções básicas).** 2 ed Jaboticabal: FUNEP, 41 p. 2003.

CROWDER, M. J.; HAND, J. **Analysis of repeated measures.** London: Chapman & Hall, 1990. 257 p.

DIGGLE, P. J. **An approach to the analysis of repeated measurements.** Biometrics, v. 44, n. 4, p. 959-971, 1988.

FERREIRA, W. L. **Análise de dados com medidas repetidas em experimento com ingestão de café.** 109 p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras, Fevereiro, 2012.

FREITAS, E. G. et al. **Modelo univariado aplicado a dados longitudinais**. Revista Brasileira Biometria, São Paulo, v. 26, n. 2, p. 93-106, 2008.

GOMES, D. P; BRINGEL, J. M. M; MORAES, M. F. H; GOMES, J. J. A; LEITE, R. M. V. B. C. **Qualidade fisiológica e sanitária de sementes de girassol produzidas na região de Timon, Maranhão**. Summa Phytopathol, Botucatu, v. 32, n. 3, 291-292 p. 2005.

HUYNH, H.; FELDT, L. S. **Conditions under which mean square ratios in repeated measurements designs have exact f-distributions**. Journal of the American Statistical Association, v. 65, n. 322, p. 1582-1589, 1970.

JOHNSON, R. A; Wichern, D. W. **Applied Multivariate Statistical Analysis**. New York. Prentice Hall, 767 p. 2002).

LIMA, G. C. **Análise de dados com medidas repetidas**. [S.l.], s.d. FZEA/USP. Apostila.

KIRK, R. E. **Experimental design: Procedures for the behavioral sciences**. Thomson Brooks/Cole Publishing Co, 1995.

MALHEIROS, E. B. **Precisão da análise de experimentos com medidas repetidas no tempo usando procedimentos do SAS**. Revista de Matemática e Estatística, São Paulo, v. 19, n. 2, p. 253-272, 2001.

MAUCHLY, J. W. **Significant test for sphericity of a normal n-variate distribution**. The Annals of Mathematical Statistics, v. 11, n. 2, p. 204-209, June 1940.

NEMEC, A. F. L. **Analysis of repeated measures and time series: an introduction with forestry examples**. Victoria: Biometric Information Handbook, 1996. 83 p.

OLIVEIRA, F. N.; TORRES, S. B.; VIEIRA, F. E. R.; PAIVA, E. P.; DUTRA, A. S. **Qualidade fisiológica de sementes de girassol avaliadas por condutividade elétrica.** e-ISSN 1983-4063 - www.agro.ufg.br/pat - Pesq. Agropec. Trop., Goiânia, v. 42, n. 3, p. 279-287, jul./set. 2012

OLIVEIRA, C. M. **Qualidade de sementes de girassol em função da posição no capítulo e da densidade de plantas.** 2013.

RENCHEER, A. C. **Methods of Multivariate Analysis.** 2 ed. Canadá, Wiley-Interscience. 727 p. 2002.

RENCHEER, A. C.; CHRISTENSEN, W. F. **Methods of Multivariate Analysis.** 3. ed. Canada: John Wiley & Sons Inc, 2012. 800 p.

ROCHA, E. B. **Modelos não-lineares para dados longitudinais provenientes de experimentos em blocos casualizados abordagem bayesiana.** 2010. 96 p. Dissertação (Mestrado em Ciências) - Escola Superior de Agricultura "Luiz de Queiroz", Piracicaba, 2010.

SILVA, S. A. A. **Análise de experimento com medidas repetidas: Uma aplicação no efeito da ingestão do café no controle de peso.** 51 p. Dissertação (Mestrado Profissional em Modelagem Matemática e Estatística Aplicada) - Universidade Vale do Rio Verde, Três Corações, Fevereiro, 2009.

SILVA, S. M. S.; ALVES, A. N.; GHEYI, H. R.; BELTRÃO, N. E. M.; SEVERINO, L. S.; SOARES, F. A. L. **Desenvolvimento e produção de duas cultivares de mamoneira sob estresse salino.** Revista Brasileira de Engenharia Agrícola e Ambiental, v. 12, n. 04, p. 335-342, 2008.

SINGER, J. M.; ROCHA, F. M. M.; NOBRE, J. S. **Análise de medidas repetidas.** In: JORNADA DE ESTATÍSTICA, 4., 2004, Maringá. Anais, Maringá: UEM, 2004. p. 120.

SOUZA, L. H. B. **Crescimento e desenvolvimento de girassol em diferentes épocas de semeadura no recôncavo da Bahia.** 100 p. Dissertação (Mestrado em Ciências Agrárias) - Universidade do recôncavo da Bahia, Cruz das Almas, abril, 2010.

TOUTENBURG, H. **Statistical analysis of designed experiments.** 2. ed. New York: Springer, 2002. 500 p. (ISSN 1431-875X).

UNGARO, M. R. G. **Potencial da cucultura do girassol com fonte de matéria-prima para o programa nacional de produção e uso de biodiesel.** In: CAMaRa, G. M.; HEIFFIG, L. S. Agronegócio de plantas oleaginosas: matérias-primas para o biodiesel. Piracicaba. Eslq, 2006. 57-80 p.

VIEIRA, F. T. P. A. **Uma abordagem multivariada em experimento silvipastoril com *Leucaena leucocephala* (Lam.) de Wit. no agreste de Pernambuco.** 70 p. Dissertação (Mestrado em Biometria) - Universidade Federal Rural de Pernambuco, Recife, 2006.

XAVIER, L. H. **Modelos univariado e multivariado para análise de medidas repetidas e verificação da acurácia do modelo univariado por meio de simulação.** 91 p. Dissertação (Mestrado em Agronomia) - Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 2000.

XAVIER, L. H.; DIAS, C. T. S. **Acurácia do modelo univariado para análise de medidas repetidas por simulação multidimensional.** Scientia Agrícola, Piracicaba, v. 58, n. 2, p. 241-250, 2001.