

Violação dos pressupostos do modelo matemático e transformação de dados

Alessandro D Lúcio; Diogo V Schwertner; Fernando M Haesbaert; Daniel Santos; Rélia R Brunes; Ana L P Ribeiro; Sidinei J Lopes

UFSM-CCR, Depto. Fitotecnia, Camobi, 97105-900 Santa Maria-RS; adlucio@ufsm.br; vs.diogo@gmail.com; fhaesbaert@gmail.com; danielsantosagro@gmail.com; relia.brunes@hotmail.com; aldpr2008@gmail.com; sjlopes@pq.cnpq.br

RESUMO

Foi verificado o atendimento às pressuposições do modelo matemático e, identificadas transformações de dados para variáveis produtivas e morfológicas de tomateiro em experimentos em túnel plástico e em campo, considerando cada data de colheita e colheitas agrupadas em duas estações de cultivo. Dois experimentos foram conduzidos em túnel plástico e um em campo, ambos no delineamento de blocos ao acaso com três repetições. Foram analisados a massa, número, comprimento e largura de frutos. Foram utilizados os testes de não aditividade de Tukey, Lilliefors, Bartlett e de Aleatoriedade para verificar respectivamente, a aditividade dos efeitos do modelo, normalidade, homogeneidade e aleatoriedade dos erros. Nos casos de violação, os dados foram transformados, e novamente submetidos à análise das pressuposições do modelo matemático. As variáveis produtivas e morfológicas de tomateiro apresentam atendimento ao pressuposto da aditividade dos efeitos. Ocorrem violações das variáveis produtivas e morfológicas às pressuposições de normalidade, homogeneidade e aleatoriedade dos erros. O agrupamento de todas as colheitas para análise proporciona atendimento das variáveis produtivas e morfológicas aos pressupostos de normalidade e homogeneidade dos erros. A transformação que possibilita maior proporção de atendimento das variáveis produtivas e morfológicas aos pressupostos de normalidade, homogeneidade e aleatoriedade dos erros é a raiz quadrada.

Palavras-chave: *Solanum lycopersicum*, agrupamento de colheitas, aditividade, normalidade, homogeneidade, aleatoriedade.

ABSTRACT

Violation of presuppositions to the mathematical model and data transformation

The adequation to the presuppositions of the mathematical model was verified and, identified data transformation for productive and morphological variables of tomato in experiments in plastic tunnel and in field, considering each harvest date and grouped harvests in two cultivation seasons. Two experiments were carried out in plastic tunnel and one in field, both in a randomized block design with three replications. We analyzed the mass, number, length and width of fruits. The tests of Tukey of non-additivity, Lilliefors, Bartlett and Run test were used to verify, respectively, the additivity of the effects of the model, normality, homogeneity and randomness of errors. In cases of violation, the data were transformed, and re-submitted to the analysis of the presuppositions of the mathematical model. The productive and morphological variables of tomato presented adequation to the presupposition of additivity of effects. Violations of the productive and morphological variables occur to the presupposition of: normality, homogeneity and randomness of errors. The grouping of all harvests for analysis provides adequation of productive and morphological variables to the presuppositions of normality and homogeneity of errors. The data transformation that provides more proportion of adequation of productive and morphological variables to the presuppositions of normality, homogeneity and randomness of errors is the square root.

Keywords: *Solanum lycopersicum*, grouping of harvests, additivity, normality, homogeneity, randomness.

(Recebido para publicação em 27 de dezembro de 2011; aceito em 9 de julho de 2012)
(Received on December 27, 2011; accepted on July 9, 2012)

No ano de 2011, o tomate (*Solanum lycopersicum*) foi a hortaliça com maior volume de produção no Brasil (4.146.466 milhões de toneladas), ocupando 66.221 mil hectares da área cultivada e com produtividade média de 62,6 t/ha. A estimativa para a safra 2012 mantém o tomateiro na liderança da produção nacional de hortaliças, com pequena redução da área de cultivo (-2%) e do volume de produção (-1,3%), compensado em parte, pelo aumento da produtividade (0,7%) (IBGE, 2012).

A correta análise e interpretação dos resultados experimentais são fundamen-

tais para a introdução de novas técnicas na agricultura, e somente são possíveis quando as pressuposições do modelo matemático são atendidas. Devido à importância econômica e comercial do tomateiro, um grande número de experimentos é realizado anualmente com a cultura. Para que ocorra a correta discriminação dos efeitos de tratamento na análise de variância, é necessário o atendimento às pressuposições de aditividade dos efeitos; os erros conjuntamente independentes, aleatórios, idêntica e normalmente distribuídos de média zero e variância comum σ^2

(Storck *et al.*, 2011).

Quando as pressuposições são violadas, a confiabilidade de todos os testes paramétricos, tais como a análise de variância, a comparação de médias e a análise de regressão ficam comprometidas, pois ocorrem alterações na probabilidade de ocorrência do erro tipo I (rejeição da hipótese de nulidade quando ela é verdadeira) e tipo II (não rejeição da hipótese de nulidade quando ela é falsa), podendo levar a falsas conclusões a respeito dos efeitos de tratamento (Steel *et al.*, 1997; Martin & Storck, 2008; Storck *et al.*, 2011).

O atendimento às pressuposições deve ser testado anteriormente à análise de variância. O teste de não aditividade de Tukey pode ser utilizado para verificar a aditividade dos efeitos (Steel *et al.*, 1997), enquanto, o teste de Lilliefors (Sprenst & Smeeton, 2007), de Bartlett (Steel *et al.*, 1997) e de Aleatoriedade (Sprenst & Smeeton, 2007) podem ser usados para verificar, respectivamente, a normalidade, a homogeneidade e a aleatoriedade dos erros. Quando os pressupostos são violados, dois caminhos são possíveis: a transformação dos dados em nova escala, de maneira que os dados obedeçam às pressuposições, ou o emprego de análises não paramétricas que possuem pressupostos mais brandos em relação aos dados do que os equivalentes testes paramétricos (Martin & Storck, 2008; Storck *et al.*, 2011).

No Rio Grande do Sul, tradicionalmente, o cultivo de tomate é em campo, sendo restrito aos meses mais quentes do ano. Entretanto, o cultivo protegido está se intensificando nessa região, por possibilitar a produção de tomate ao longo do ano todo, com incremento na produtividade e qualidade de frutos, maior precocidade de produção, facilidade no controle de pragas e doenças e economia da água de irrigação (Cermeño, 1990).

Características do ambiente protegido e da cultura do tomate podem contribuir para uma maior variabilidade experimental e dificultar o atendimento às pressuposições do modelo matemático. Dentre as características do ambiente protegido, podem-se citar a proximidade das linhas de cultivo com as laterais do túnel e o manejo cultural intensivo. Já, como características do tomateiro, podem-se citar as sucessivas colheitas realizadas e a ausência de frutos aptos à colheita em momentos específicos (Lúcio *et al.*, 2010a, 2010b).

Trabalhos visando identificar violações das pressuposições do modelo matemático e consequências sobre as análises dos resultados experimentais foram realizados em: milho (Marques *et al.*, 2000), sorgo (Lopes *et al.*, 2009), alface (Zanardo *et al.*, 2010), abobrinha italiana (Carpes *et al.*, 2010) e tomate (Lúcio *et al.*, 2010a). Em trabalhos com abobrinha italiana (Couto *et al.*, 2009), com brócolis e com pimentão (Lúcio *et*

al., 2010b, 2011) além da violação aos pressupostos do modelo matemático, foram investigadas transformações de dados, utilizando a metodologia Box-Cox (Box & Cox, 1964). Os resultados para hortaliças de múltiplas colheitas têm indicado violação das pressuposições de normalidade e homogeneidade dos erros, bem como, dificuldades em obter transformações eficientes dos dados, sendo necessárias transformações diferentes entre variáveis de uma mesma espécie, para a mesma variável em diferentes espécies e, para a mesma variável e espécie em diferentes épocas de cultivo.

Para hortaliças, a maioria dos trabalhos limita-se a verificar o atendimento aos pressupostos da normalidade e homogeneidade dos erros das variáveis massa e número de frutos. A ausência de informações a respeito do atendimento às pressuposições de aditividade dos efeitos e aleatoriedade dos erros em experimentos com hortaliças, associados à ausência de estudos com tomate, visando identificar o atendimento às pressuposições do modelo matemático e a necessidade de transformação de dados, leva ao objetivo deste trabalho, que foi: verificar o atendimento às pressuposições do modelo matemático e, identificar transformações de dados para variáveis produtivas e morfológicas de tomateiro em experimentos em túnel plástico e em campo, considerando cada data de colheita e colheitas agrupadas em duas estações de cultivo.

MATERIAL E MÉTODOS

Foram conduzidos três experimentos com tomate tipo salada, híbrido Grandeur, na Universidade Federal de Santa Maria, em Santa Maria (RS) (29°43'S, 53°43'W, altitude 95 m). O clima da região, segundo a classificação de Köppen (Moreno, 1961) é do tipo Cfa (subtropical úmido sem estação seca definida e com verões quentes) e o solo classificado como Argissolo Vermelho Distrófico arênico (Sistema Brasileiro de Classificação de solos, 2006).

Dois experimentos foram realizados simultaneamente na estação primavera-verão (P-V) de 16/08/10 a 27/01/11, e o terceiro na estação outono-inverno

(O-I) de 03/03/11 a 30/07/11. Na P-V, um experimento foi em túnel plástico e outro em campo e no O-I, o experimento foi em túnel. O túnel possui 3,5 m de altura na parte central, 25 m de comprimento por 4 m de largura, com dimensão útil de 19,2 m de comprimento e 3,6 m de largura, coberto com filme de polietileno de baixa densidade (PEBD) de 100 micras, com aditivo anti-UV, e orientação Norte-Sul.

Nos três experimentos, as mudas foram transplantadas com quatro folhas verdadeiras e dispostas em três linhas (camalhões sem uso de mulching) com 15 cm de altura e 40 cm de largura e, irrigação por gotejamento. Utilizou-se espaçamento de 0,8 m entre plantas e 1,2 m entre linhas, num total de 24 plantas por linha de cultivo.

A adubação no transplante foi realizada com: 65 kg ha⁻¹ de N, 230 kg ha⁻¹ de P₂O₅ e 65 kg ha⁻¹ de K₂O. A adubação de amontoa foi realizada aos 21 dias após o transplante com: 35 kg ha⁻¹ de N e 35 kg ha⁻¹ de K₂O, e as adubações de cobertura com 30 kg ha⁻¹ de N e 30 kg ha⁻¹ de K₂O, aplicados em intervalos de 15 dias, iniciando após a adubação de amontoa, num total de sete aplicações durante o ciclo. Todas as adubações, bem como a calagem, foram realizadas conforme resultados da análise de solo por linha de cultivo (Sociedade Brasileira de Ciência do Solo, 2004). As plantas foram conduzidas em haste única e, todos os demais tratamentos culturais foram realizados conforme as recomendações para a cultura (Filgueira, 2008).

A parcela foi constituída de duas plantas na direção da linha de cultivo. Foi utilizado o delineamento blocos ao acaso, com três repetições e 12 tratamentos. Os tratamentos foram a combinação de dois intervalos de aplicação de Acrescent Solus® (após todas as colheitas de frutos e em colheitas alternadas) com a adubação mineral recomendada em cobertura mais as doses de 1, 2, 3 e 4 L ha⁻¹ de Acrescent Solus® aplicados no solo, e: (T₁) adubação mineral em cobertura (sem Acrescent Solus®); (T₂) adubação mineral em cobertura mais 50 L ha⁻¹ de Acrescent Solus® aplicados aos 30 e 60 dias após o transplante; (T₃) substituição da adubação mineral em cobertura pela dose de 100 L ha⁻¹

de Acrescent Solus® aplicados a cada 15 dias e; (T_4) adubação mineral em cobertura mais 0,5 L ha⁻¹ de Acrescent Solus® aplicados após todas as colheitas.

Na estação primavera-verão, foram realizadas dez colheitas (de 11/11/10 a 27/01/11) e, no outono-inverno, seis (de 02/06/11 a 30/07/11). Em cada colheita, foram avaliados a massa de frutos por planta em balança digital com precisão de 1 g; o número de frutos por planta; o comprimento e a largura média de frutos, mensurados com paquímetro com precisão de 1 mm.

O atendimento de cada variável aos pressupostos do modelo matemático foi testado em cada data de colheita e em colheitas agrupadas. Para isso, os grupos formados seguiram a sequência da divisão das colheitas obtidas pelos seus múltiplos nos diferentes experimentos. Os agrupamentos testados foram os seguintes: agrupamento duas em duas (1^a+2^a ; 3^a+4^a ; ...; 9^a+10^a), agrupamento cinco em cinco ($1^a+2^a+3^a+4^a+5^a$; $6^a+7^a+8^a+9^a+10^a$) e agrupamento total ($1^a+2^a+3^a$; ...; $8^a+9^a+10^a$) na primavera-verão; e, agrupamento duas em duas (1^a+2^a ; 3^a+4^a ; 5^a+6^a), agrupamento três em três ($1^a+2^a+3^a$; $4^a+5^a+6^a$) e agrupamento total (1^a+2^a ; ...; 5^a+6^a), no outono-inverno.

Para testar o atendimento às pressuposições do modelo matemático e obter o erro experimental foi considerado o modelo matemático do delineamento blocos ao acaso: $y_{ij} = m + t_i + b_j + e_{ij}$, com $i = 12$ tratamentos e, $j = 3$ repetições, onde: m é uma constante associada a todas as observações; t_i é o efeito do tratamento; b_j é o efeito do bloco; e_{ij} é o erro experimental associado à observação y_{ij} (Storck *et al.*, 2011). Para testar a aditividade dos efeitos do modelo foi utilizado o teste de não aditividade de Tukey (Steel *et al.*, 1997), enquanto que, para testar a normalidade, homogeneidade e aleatoriedade dos erros foram utilizados os testes de Lilliefors (Sprent & Smeeton, 2007), Bartlett (Steel *et al.*, 1997) e de Aleatoriedade (Sprent & Smeeton, 2007), respectivamente.

Para os casos em que ocorreu alguma violação dos pressupostos do modelo matemático, foram realizadas transformações de dados, e os mesmos foram novamente submetidos à análise

do atendimento das pressuposições do modelo matemático para verificar se a transformação foi eficiente em adequar a variável ao pressuposto violado. Foram testadas transformações já utilizadas em outros estudos e presentes na literatura: raiz quadrada ($(Y_{ij})^{0.5}$) (Storck *et al.*, 2011), raiz quarta ($(Y_{ij})^{0.25}$) (Couto *et al.*, 2009) e suas inversas ($(Y_{ij})^{-0.5}$ e $(Y_{ij})^{-0.25}$) (Lúcio *et al.*, 2010b; 2011); inversa ou recíproca ($(Y_{ij})^{-1}$) (Storck *et al.*, 2011). Nestas transformações, quando havia no banco de dados valores nulos adicionou-se uma constante ($c = 0,5$) somada à variável (Storck *et al.*, 2011). Foi também testada a transformação logarítmica ($\log_{10}(Y_{ij})$) e sua inversa ($\log_{10}(Y_{ij})^{-1}$), adicionando-se a constante ($c = 1$) para valores nulos (Lúcio *et al.*, 2010b; Storck *et al.*, 2011). Todos os testes foram aplicados em nível de 5% de probabilidade de erro, utilizando planilha eletrônica no aplicativo Office Excel® elaborada por Martin & Storck (2008).

RESULTADOS E DISCUSSÃO

Houve 100% de atendimento ao pressuposto da aditividade dos efeitos do modelo para todas as variáveis analisadas em colheitas individuais e agrupadas, no ambiente túnel plástico e em campo nas duas estações de cultivo (Tabelas 1, 2 e 3). Alta porcentagem de atendimento dessa pressuposição também foi obtida por Marques *et al.* (2000) em ensaios de competição de cultivares de milho, nos quais verificaram a violação desse pressuposto em apenas 8,75%, dos 307 experimentos avaliados.

Foi verificado comportamento diferenciado no atendimento aos pressupostos da normalidade, homogeneidade e aleatoriedade dos erros entre as variáveis estudadas, ambientes e estações de cultivo (Tabelas 1, 2 e 3). As variáveis massa (MF) e número de frutos (NF), apresentaram, em cada tipo de colheita, porcentagem de atendimento aos pressupostos da normalidade e homogeneidade dos erros maior, ou no mínimo igual, ao das variáveis comprimento (COM) e largura (LAR) média de frutos. A relação direta entre a massa e o NF de tomate ($r = 0,88$), pode ser a explicação para o similar comportamento dessas

variáveis, assim como, observado em abobrinha italiana (Couto *et al.*, 2009). O menor percentual de atendimento, a normalidade e homogeneidade dos erros das variáveis COM e LAR em relação a massa e NF, pode ser explicado, pelo fato daquelas apresentarem maior variabilidade, causando a assimetria, que resulta também em heterogeneidade dos erros (Martin & Storck, 2008).

Para a aleatoriedade dos erros, observou-se comportamento inverso, com maior ou igual porcentagem de adequação à pressuposição das variáveis COM e LAR de frutos, em relação a massa e NF (Tabelas 1, 2 e 3). Variáveis produtivas como MF e NF, caracterizam-se por apresentar baixa herdabilidade (16,54 e 22,09%, respectivamente) e, portanto, sofrem maior influência de fatores não controlados adequadamente pelo delineamento experimental, responsáveis pela falta de aleatoriedade dos erros (Marques *et al.*, 2000; Martin & Storck, 2008). As variáveis COM e LAR, por apresentarem maior herdabilidade (63,14 e 48,30%, respectivamente), são menos influenciadas pela heterogeneidade da área experimental, possuindo, portanto, maior atendimento ao pressuposto da aleatoriedade dos erros.

De maneira geral, na primavera-verão, o cultivo em campo ocasionou maiores porcentagens de atendimento aos pressupostos da normalidade, homogeneidade e aleatoriedade dos erros em relação ao cultivo em túnel plástico (Tabelas 1 e 2). Isso ocorre porque o túnel é um ambiente de cultivo mais homogêneo do que em campo, com maior amplitude térmica e de umidade do ar; sombreamento ocasionado pela estrutura; as duas linhas laterais do túnel ficam em condições diferentes da linha central quanto à temperatura, umidade do ar e do solo, ocasionando aumento da variabilidade experimental pela heterogeneidade entre as linhas de cultivo e, em consequência, redução no atendimento às pressuposições do modelo matemático (Lúcio *et al.*, 2003, 2008; Carpes *et al.*, 2008).

Menores porcentagens de atendimento às pressuposições de normalidade, homogeneidade e aleatoriedade dos erros, foram observadas de maneira

Tabela 1. Porcentagem de atendimento aos pressupostos do modelo matemático das variáveis massa de frutos (g planta⁻¹) (MF), número de frutos por planta (NF), comprimento médio de frutos (COM, mm) e largura média de frutos (LAR, mm) de tomate, em colheitas individuais e agrupadas na primavera-verão (percentage of adequation to the presuppositions of the mathematical model of variables mass of fruit (g plant⁻¹) (MF), number of fruits per plant (NF), average length of fruits (COM, mm) and average width of fruits (LAR, mm) of tomato, in individual and grouped harvests in spring-summer). Santa Maria, UFSM, 2011.

Primavera-verão em túnel				
Tipo de colheita	Variáveis			
	MF	NF	COM	LAR
Aditividade (%)				
Individuais	100	100	100	100
Agrupadas 2 em 2	100	100	100	100
Agrupadas 5 em 5	100	100	100	100
Total	100	100	100	100
Normalidade (%)				
Individuais	100	100	90	90
Agrupadas 2 em 2	100	100	100	80
Agrupadas 5 em 5	100	100	100	100
Total	100	100	100	100
Homogeneidade (%)				
Individuais	80	70	50	40
Agrupadas 2 em 2	80	100	40	80
Agrupadas 5 em 5	100	50	100	100
Total	100	100	100	100
Aleatoriedade (%)				
Individuais	10	0	10	20
Agrupadas 2 em 2	0	0	0	20
Agrupadas 5 em 5	0	0	0	0
Total	0	0	0	0

geral no cultivo em túnel plástico no outono-inverno em relação ao cultivo na primavera-verão (Tabelas 2 e 3), também relatado por Lúcio *et al.* (2008), Carpes *et al.* (2008) e Couto *et al.* (2009), em abobrinha italiana. Essas diferenças podem ser atribuídas às condições ambientais desfavoráveis ao desenvolvimento do tomateiro no outono-inverno, como temperatura do ar inferior a ideal (21-24°C) e, inferior a mínima (10°C) para a cultura (Filgueira, 2008). Também, nesta época (maio, junho e julho) a radiação solar atinge valores insuficientes para permitir o crescimento normal das plantas no município de Santa Maria, prejudicando o processo fotossintético (Estefanel *et al.*, 1998), ocasionando abortamento de flores, aumento da frequência de valores zero de produção no banco de dados e inflacionando a variância experimental

(Lúcio *et al.*, 2010b).

Houve 100% de atendimento ao pressuposto da normalidade dos erros no cultivo em campo (Tabela 2), enquanto no túnel, houve violação dessa pressuposição em colheitas individuais para as variáveis comprimento (COM) e largura (LAR) de frutos e no agrupamento de duas colheitas, na variável LAR na primavera-verão (P-V) (Tabela 1). No outono-inverno (O-I), houve violação desse pressuposto em todas as variáveis em colheitas individuais e apenas para a variável COM no agrupamento de duas colheitas (Tabela 3). A partir do agrupamento da produção em duas metades iguais (agrupamento de cinco colheitas na P-V e de três colheitas no O-I) verifica-se ausência de violação ao pressuposto da normalidade dos erros, fundamental para a validação dos testes de hipóteses.

A ausência de normalidade dos erros geralmente causa um aumento no verdadeiro nível de significância do teste de hipótese, resultando em maior probabilidade de erro tipo I do que o especificado pelo nível de significância adotado, além da falta de normalidade ser uma das possíveis causas da heterogeneidade regular dos erros, quando existe uma relação entre a média e a variância dos vários tratamentos (Steel *et al.*, 1997; Storck *et al.*, 2000; Martin & Storck, 2008).

Verifica-se nas tabelas 1, 2 e 3, uma proporção menor de atendimento ao pressuposto da homogeneidade dos erros em relação ao observado para a normalidade dos erros no mesmo ambiente, variável e agrupamento de colheita, o que indica não ser a falta de normalidade dos dados a principal causa da heterocedasticidade. A falta de homogeneidade dos erros pode ser atribuída no mínimo à natureza e características dos tratamentos (Storck *et al.*, 2000), aos manejos aplicados (Carpes *et al.*, 2008), à escolha incorreta da unidade experimental (Lúcio *et al.*, 2010a), à heterogeneidade da área experimental, às múltiplas colheitas realizadas, ao excesso de valores zero de produção (Couto *et al.*, 2009; Lúcio *et al.*, 2010a, 2010b) e, à variável e a forma de avaliação (Lopes *et al.*, 2009). Somente com o agrupamento de todas as colheitas realizadas no túnel plástico foi possível obter 100% de homogeneidade dos erros para todas as variáveis analisadas, enquanto no cultivo em campo, já a partir do agrupamento de cinco colheitas, obtém-se o mesmo nível de atendimento a essa pressuposição.

Pequenas porcentagens de atendimento à pressuposição da aleatoriedade dos erros foi obtida para a ampla maioria das variáveis observadas, independente do tipo de colheita, época e ambiente de cultivo (Tabelas 1, 2 e 3), exceção apenas à variável comprimento de frutos, na primavera-verão, no total da produção e ambiente túnel (Tabela 2), que não apresentou violação dessa pressuposição. Não atendimento a essa pressuposição do modelo matemático também foi verificada em 9,88% de 307 ensaios de milho realizados no RS entre 1993 e 1996, sendo a consequên-

Tabela 2. Porcentagem de atendimento aos pressupostos do modelo matemático das variáveis massa de frutos (g planta⁻¹) (MF), número de frutos por planta (NF), comprimento médio de frutos (COM, mm) e largura média de frutos (LAR, mm) de tomate, em colheitas individuais e agrupadas na primavera-verão (percentage of adequation to the presuppositions of the mathematical model of variables mass of fruit (g plant⁻¹) (MF), number of fruits per plant (NF), average length of fruits (COM, mm) and average width of fruits (LAR, mm) of tomato, in individual and grouped harvests in spring-summer). Santa Maria, UFSM, 2011.

Primavera-verão em campo				
Tipo de colheita	Variáveis			
	MF	NF	COM	LAR
Aditividade (%)				
Individuais	100	100	100	100
Agrupadas 2 em 2	100	100	100	100
Agrupadas 5 em 5	100	100	100	100
Total	100	100	100	100
Normalidade (%)				
Individuais	100	100	100	100
Agrupadas 2 em 2	100	100	100	100
Agrupadas 5 em 5	100	100	100	100
Total	100	100	100	100
Homogeneidade (%)				
Individuais	70	60	40	40
Agrupadas 2 em 2	100	100	40	80
Agrupadas 5 em 5	100	100	100	100
Total	100	100	100	100
Aleatoriedade (%)				
Individuais	0	10	20	50
Agrupadas 2 em 2	20	20	40	20
Agrupadas 5 em 5	0	0	50	50
Total	0	0	100	0

cia direta dessa violação o aumento da variância experimental e o erro tipo II (Marques *et al.*, 2000).

A falta de aleatoriedade dos erros não era um resultado esperado, pois em princípio, a aleatoriedade é assegurada pela casualização dos tratamentos na área experimental (Storck *et al.*, 2011; Martin & Storck, 2008). Conforme os autores, a falta de aleatoriedade pode indicar que os blocos foram dispostos de forma inadequada à variabilidade existente na área experimental. No entanto, trabalhos têm demonstrado a necessidade do uso de blocos no sentido longitudinal do ambiente protegido (estufas e túneis) para controlar a heterogeneidade entre as linhas de cultivo ocasionadas pela proximidade com as laterais (Lorentz *et al.*, 2005; Carpes *et al.*, 2008; Lúcio *et al.*, 2008).

A não aleatoriedade dos erros pode

ser atribuída à elevada variabilidade existente dentro das linhas de cultivo, ocasionada possivelmente por diferenças de fertilidade do solo, heterogeneidade de distribuição de água pelo sistema de irrigação por gotejamento e ao pequeno tamanho de parcela utilizado (Carpes *et al.*, 2008) pois, observando os mapas que contêm a estimativa dos erros de cada parcela foi possível identificar que: a) em determinados blocos, os erros são negativos e poucos são positivos (ou vice-versa); b) os erros positivos ou negativos estão em muitos casos agrupados em posições intermediárias ou nas extremidades dos blocos e; c) em pelo menos um dos blocos, ocorreu concentração dos erros positivos e/ou negativos em posição intermediária ou na sua extremidade, concordando com resultados apresentados por Marques *et al.* (2000) em situações de falta de aleatoriedade

dos erros, em ensaios de milho. São necessários portanto, novos estudos de simulação de tamanho de parcela para tomateiro, tendo por objetivo identificar um tamanho de parcela experimental que possibilite maior atendimento ao pressuposto da aleatoriedade dos erros, como já realizado para o pressuposto da homogeneidade dos erros (Lúcio *et al.*, 2010a).

O agrupamento das sucessivas colheitas demonstrou ser uma técnica viável para aumentar as porcentagens de atendimento aos pressupostos da normalidade e homogeneidade dos erros (Tabelas 1, 2 e 3) principalmente a partir do agrupamento de cinco colheitas na primavera-verão e total (seis colheitas) no outono-inverno, concordando com os resultados obtidos em experimento em branco com tomateiro, em que para parcelas de duas plantas, observaram-se colheitas homogêneas a partir do agrupamento de seis colheitas (Lúcio *et al.*, 2010a), enquanto que, para abobrinha italiana, já a partir do agrupamento de três colheitas (Carpes *et al.*, 2010). Porém, o agrupamento de colheitas não é eficiente em proporcionar maiores porcentagens de atendimento ao pressuposto da aleatoriedade dos erros, pois mesmo com o agrupamento do total de colheitas, observa-se que apenas há atendimento dessa pressuposição para a variável comprimento de frutos no cultivo em campo (Tabela 2).

O agrupamento de colheitas permitiu suavizar os valores extremos de produção observados em colheitas individuais, principalmente, pelo preenchimento das lacunas formadas pelas parcelas não colhidas (produção nula na colheita), responsável pelo não atendimento às pressuposições de normalidade e homogeneidade dos erros (Lúcio *et al.*, 2010b). Com o agrupamento do total de colheitas observou-se aumento da média de todas as variáveis e; redução no desvio padrão das variáveis comprimento e largura de frutos; ausência de valores zero de produção, e assim, maior uniformidade entre as diferentes repetições de um mesmo tratamento, o que proporciona melhor atendimento às pressuposições de normalidade e homogeneidade dos erros (Carpes *et al.*, 2010; Lúcio *et al.*, 2010a, 2010b). A falta de

Tabela 3. Porcentagem de atendimento aos pressupostos do modelo matemático das variáveis massa de frutos (g planta⁻¹) (MF), número de frutos por planta (NF), comprimento médio de frutos (COM, mm) e largura média de frutos (LAR, mm) de tomate, em colheitas individuais e agrupadas no outono-inverno (percentage of adequation to the presuppositions of the mathematical model of variables mass of fruit (g plant⁻¹) (MF), number of fruits per plant (NF), average length of fruits (COM, mm) and average width of fruits (LAR, mm) of tomato, in individual and grouped harvests in autumn-winter). Santa Maria, UFSM, 2011.

Outono-inverno em túnel				
Tipo de colheita	Variáveis			
	MF	NF	COM	LAR
Aditividade (%)				
Individuais	100,0	100,0	100,0	100,0
Agrupadas 2 em 2	100,0	100,0	100,0	100,0
Agrupadas 3 em 3	100,0	100,0	100,0	100,0
Total	100,0	100,0	100,0	100,0
Normalidade (%)				
Individuais	83,3	83,3	66,7	66,7
Agrupadas 2 em 2	100,0	100,0	66,7	100,0
Agrupadas 3 em 3	100,0	100,0	100,0	100,0
Total	100,0	100,0	100,0	100,0
Homogeneidade (%)				
Individuais	83,3	66,7	0,0	33,3
Agrupadas 2 em 2	100,0	33,3	33,3	0,0
Agrupadas 3 em 3	100,0	50,0	0,0	0,0
Total	100,0	100,0	100,0	100,0
Aleatoriedade (%)				
Individuais	0,0	16,7	16,7	0,0
Agrupadas 2 em 2	0,0	0,0	0,0	0,0
Agrupadas 3 em 3	0,0	0,0	0,0	0,0
Total	0,0	0,0	0,0	0,0

aleatoriedade dos erros não foi superada pelo agrupamento das colheitas porque está intimamente condicionada à elevada heterogeneidade da área experimental (Marques *et al.*, 2000).

Uma alternativa à análise de dados com violação de pressupostos do modelo matemático é o uso de análises não paramétricas, pois não exigem normalidade e homogeneidade dos erros (Steel *et al.*, 1997; Sprent & Smeeton, 2007). Contudo, esses também exigem a aleatoriedade dos erros, não observada nesse estudo. Embora exista um grande número de testes não paramétricos disponíveis, esses são de aplicação restrita e de resultados pouco satisfatórios na maioria dos casos, sendo a prática mais comum, entre os pesquisadores, a transformação de dados para adequar a variável às pressuposições violadas (Storck *et al.*, 2011).

Na figura 1 pode ser observada a porcentagem de eficiência das transformações testadas, conforme a variável, ambiente de cultivo, tipo de colheita e pressuposição violada. São apresentados resultados para as situações em que foi possível obter pelo menos uma transformação eficiente para os casos de violação observados nas tabelas 1, 2 e 3.

A transformação raiz quadrada $((Y+0,5)^{0,5})$ apresentou 100% de eficiência em adequar a massa e o número de frutos ao pressuposto da normalidade dos erros, sem apresentar eficiência para as variáveis comprimento e largura de frutos (Figura 1c). Porém, dentre todos os casos de violação desse pressuposto (Tabelas 1, 2 e 3), essa transformação foi eficiente apenas em colheitas individuais, no cultivo em túnel plástico no outono-inverno. A difícil detecção de transformações adequadas

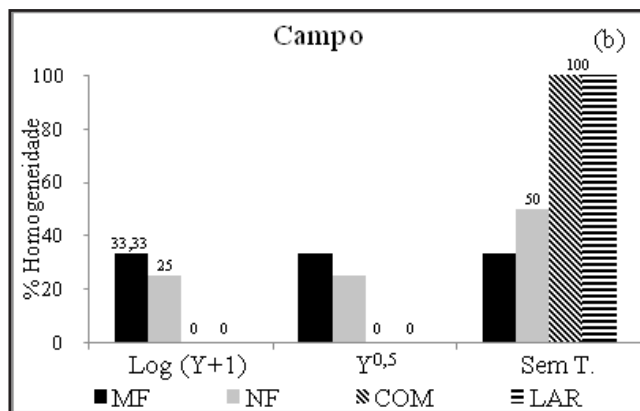
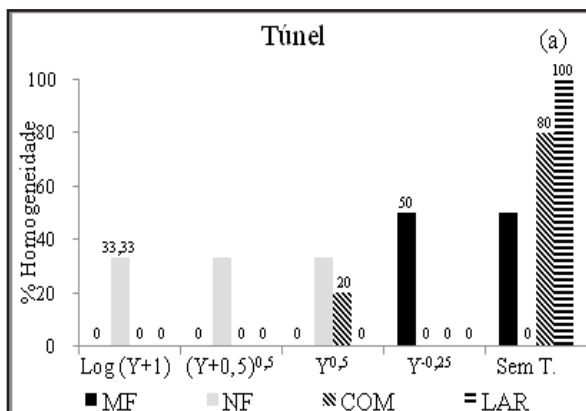
para superar a falta de normalidade dos erros indica que os desvios observados são severos e de difícil contorno, também relatado por Lúcio *et al.* (2010b), em brócolis, em que, mesmo após a transformação dos dados não houve aderência dos erros à normalidade.

A transformação logarítmica (Log (Y+1)) foi eficiente em 33,33% dos casos de violação do pressuposto da homogeneidade dos erros em colheitas individuais do número de frutos (NF) no cultivo em túnel (Figura 1a) e, em 33,33% e 25% dos casos de violação das variáveis massa e NF, respectivamente, em campo (Figura 1b), ambos na primavera-verão. Storck *et al.*, (2011) e Martin & Storck, (2008) enfatizam que a transformação logarítmica é geralmente eficiente nos casos de heterocedasticidade regular, quando a variância é proporcional à média do tratamento. A transformação logarítmica também foi identificada, pelo procedimento Box-Cox como a mais adequada para cabeças e ramificações comerciais e não comerciais de brócolis, minimizando a variabilidade, mas sem assegurar a homogeneidade dos erros (Lúcio *et al.*, 2010b).

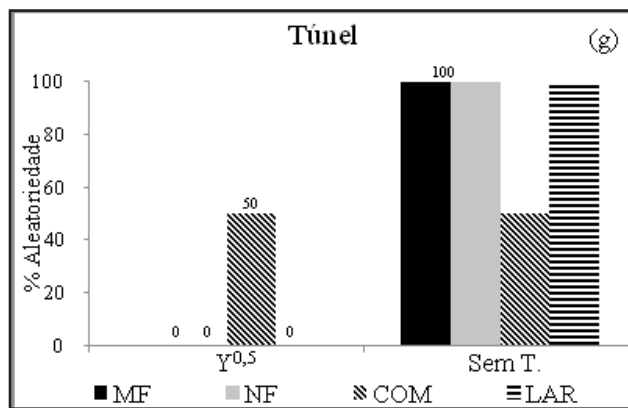
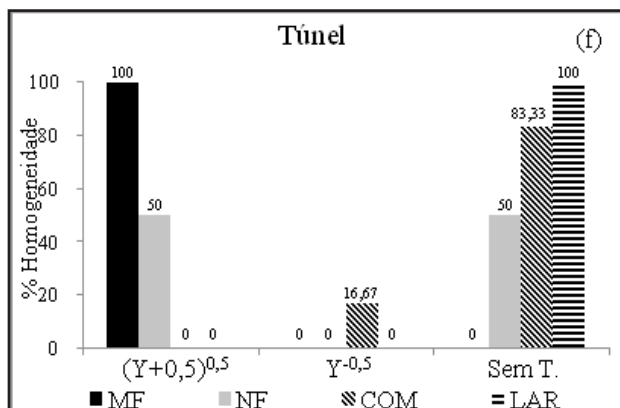
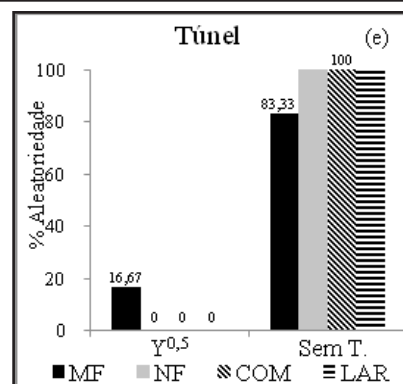
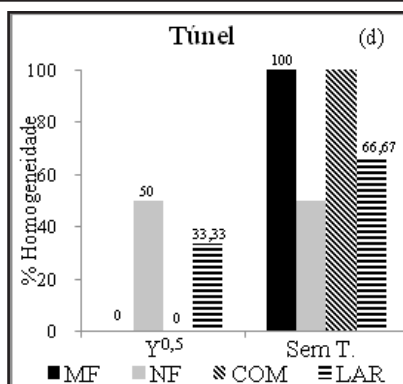
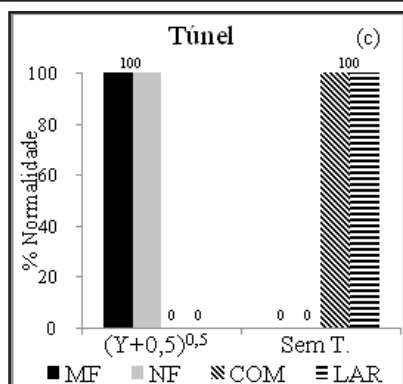
As transformações inversas da raiz quarta $(Y^{0,25})$ e raiz quadrada $((Y+0,5)^{0,5})$ apresentaram 33,33 e 50% de eficiência respectivamente em adequar as variáveis número (NF) e massa de frutos ao pressuposto da homogeneidade dos erros, na primavera-verão, no cultivo em túnel em colheitas individuais (Figura 1a). Já no cultivo em campo, para a mesma situação (Figura 1b), essas transformações não foram eficientes. A transformação $((Y+0,5)^{0,5})$ também foi eficiente no outono-inverno, em colheitas individuais em 100 e 50% dos casos de violação das variáveis massa e NF (Figura 1f). A transformação inversa da raiz quarta também foi identificada como a mais adequada para massa e NF de pimentão (Lúcio *et al.*, 2011), cabeças e ramificações comerciais e não comerciais de brócolis, no outono (Lúcio *et al.*, 2010b), minimizando a variabilidade, mas sem assegurar a homogeneidade dos erros.

A transformação raiz quadrada $(Y^{0,5})$ apresentou 33,33 e 20% de eficiência em adequar as variáveis NF e com-

Primavera-Verão



Outono-Inverno



Log (Y+1)= transformação logarítmica acrescida da constante c=1 (logarithmic transformation plus constant); $(Y+0,5)^{0,5}$ = raiz quadrada acrescida da constante c=0,5 (square root plus constant); $Y^{0,5}$ = raiz quadrada (square root); $Y^{-0,25}$ = inversa da raiz quarta (inverse fourth root); $Y^{-0,5}$ = inversa da raiz quadrada (inverse square root).

Figura 1. Eficiência percentual da transformação de dados em provocar o atendimento das variáveis massa (MF) e número de frutos (NF) por planta, comprimento (COM) e largura (LAR) média de frutos de tomate ao pressuposto violado e porcentagem de casos sem transformação eficiente (Sem T.) ('a', 'b', 'c', 'e', 'f')= colheitas individuais; 'd'= colheitas agrupadas duas em duas; 'g'= colheitas agrupadas três em três) (percentage efficiency of data transformation to adequate the variables fruit mass (MF) and number of fruits (NF) per plant, length (COM) and width (LAR) average of fruits of tomato to the violated presuppositions and percentage of cases without efficient transformation (Sem T.) ('a', 'b', 'c', 'e', 'f')= individual harvests; 'd'= harvests grouped two in two; 'g'= harvests grouped three in three)). Santa Maria, UFSM, 2011.

primeto de frutos ao pressuposto da homogeneidade dos erros em colheitas individuais no ambiente túnel (Figura 1a); e, 33,33 e 25% de eficiência para

as variáveis massa e NF, também em colheitas individuais em campo (figura 1b), ambos na primavera-verão. No outono-inverno, a mesma transformação

foi eficiente em 50 e 33% dos casos de violação da homogeneidade dos erros das variáveis NF e largura de frutos (Figura 1d) no agrupamento de colhei-

tas duas em duas. A transformação raiz quadrada é geralmente eficiente quando os dados possuem poucos valores extremos, como valores muito baixos de produção observados em determinadas repetições de um tratamento pela ausência de frutos aptos a colheita (Martin & Storck, 2008; Storck *et al.*, 2011). Ainda, a transformação inversa da raiz quadrada ($Y^{0.5}$)⁻¹ foi eficiente em 16,67% dos casos de violação desta pressuposição para o comprimento de frutos, apenas no outono-inverno, em colheitas individuais (Figura 1f), também identificada por Lúcio *et al.*, (2010b) como a transformação mais adequada para ramificações comerciais de brócolis no outono.

Para superar a falta de aleatoriedade dos erros, foi identificada a transformação raiz quadrada ($Y^{0.5}$), que permitiu a adequação de 16,67 e 50% dos casos de violação dessa pressuposição das variáveis massa e comprimento de frutos, respectivamente, em colheitas individuais (Figura 1e) e agrupadas três em três (Figura 1g) no outono-inverno. Não era esperado identificar uma transformação de dados que permitisse superar a falta de aleatoriedade dos erros, pois conforme Martin & Storck, (2008) não há solução aparente para a não aleatoriedade dos erros, tipo de transformação de dados ou uso de testes não paramétricos.

Dessa forma, fica evidente que a transformação adequada depende da variável observada, estação de cultivo, ambiente experimental, agrupamento de colheita e pressuposição violada. Os resultados apresentados são uma primeira indicação do tipo de transformação a ser testada quando os dados apresentam violação dos pressupostos do modelo matemático, podendo acelerar o processo de identificação da transformação adequada.

Foi possível identificar transformação eficiente de dados em apenas 22,78% das 79 situações de violação às pressuposições de normalidade, homogeneidade e aleatoriedade dos erros observados nas Tabelas 1, 2 e 3. Mesmo dentro desses 22,78% de situações de violação, que apresentaram pelo menos uma transformação adequada, a eficiência das transformações é na sua

maioria inferior a 50% do número de casos de violação observados (Figura 1), enfatizando a dificuldade de identificar transformações de dados eficientes em adequar a variável ao pressuposto violado em experimentos com hortaliças de múltiplas colheitas (Couto *et al.*, 2009; Lúcio *et al.*, 2010b, 2011).

As variáveis produtivas e morfológicas de tomateiro apresentam atendimento ao pressuposto da aditividade dos efeitos. Ocorrem violações das variáveis produtivas e morfológicas às pressuposições de normalidade, homogeneidade e aleatoriedade dos erros. O agrupamento de todas as colheitas para análise proporciona atendimento das variáveis produtivas e morfológicas aos pressupostos de normalidade e homogeneidade dos erros. A transformação que possibilita maior proporção de atendimento das variáveis produtivas e morfológicas aos pressupostos de normalidade, homogeneidade e aleatoriedade dos erros é a raiz quadrada.

AGRADECIMENTOS

À CAPES, pela concessão de bolsa de mestrado e ao CNPq e FAPERGS pelo auxílio financeiro para a realização do trabalho e pela concessão de bolsa de produtividade em pesquisa ao primeiro e ao último autor.

REFERÊNCIAS

- BOX GEP; COX DR. 1964. An analysis of transformations. *Journal of the Royal Society* 26: 211-252.
- CARPES RH; LÚCIO AD; LOPES SJ; BENZ V; HAESBAERT F; SANTOS D. 2010. Variabilidade produtiva e agrupamentos de colheitas de abobrinha italiana cultivada em ambiente protegido. *Ciência Rural* 40: 294-301.
- CARPES RH; LÚCIO AD; STORCK L; LOPES SJ; ZANARDO B; PALUDO AL. 2008. Ausência de frutos colhidos e suas interferências na variabilidade da fitomassa de frutos de abobrinha italiana cultivada em diferentes sistemas de irrigação. *Revista Ceres* 55: 590-595.
- CERMEÑO ZS. 1990. *Estufas – instalações e manejo*. Lisboa: Litexa. 355p.
- COUTO MRM; LÚCIO AD; LOPES SJ; CARPES RH. 2009. Transformação de dados em experimentos com abobrinha italiana em ambiente protegido. *Ciência Rural* 39: 1701-1707.
- ESTEFANEL V; BURIOL GA; ANDRIOLO JL; LIMA CP; LUZZI N. 1998. Disponibilidade de radiação solar nos meses de inverno para o cultivo do tomateiro (*Lycopersicon esculentum* Mill.) na região de Santa Maria, RS. *Ciência Rural* 28: 553-559.
- FILGUEIRA FAR. 2008. *Novo manual de Olericultura: Agrotecnologia moderna na produção e comercialização de hortaliças*. Viçosa: UFV. 421p.
- IBGE. 2012, 9 de abril. *Levantamento sistemático da produção agrícola*. Disponível em http://www.ibge.gov.br/home/estatistica/indicadores/agropecuaria/lspa/lspa_201202.pdf
- LOPES SJ; BRUM B; STORCK L; LÚCIO AD; SILVEIRA TR; TOEBE M. 2009. Espaçamento entre plantas de sorgo granífero: produtividade de grãos e qualificação do modelo matemático. *Ciência Rural* 39: 649-656.
- LORENTZ LH; LÚCIO AD; BOLIGON AA; LOPES SJ; STORCK L. 2005. Variabilidade da produção de frutos de pimentão em estufa plástica. *Ciência Rural* 35: 316-323.
- LÚCIO AD; CARPES RH; STORCK L; LOPES SJ; LORENTZ LH; PALUDO AL. 2008. Variância e média da massa de frutos de abobrinha-italiana em múltiplas colheitas. *Horticultura Brasileira* 26: 335-341.
- LÚCIO AD; CARPES RH; STORCK L; ZANARDO B; TOEBE M; PUHL OJ; SANTOS JRA. 2010a. Agrupamento de colheitas de tomate e estimativas do tamanho de parcela em cultivo protegido. *Horticultura Brasileira* 28: 190-196.
- LÚCIO AD; COUTO MRM; TREVISAN JN; MARTINS GAK; LOPES SJ. 2010b. Excesso de zeros nas variáveis observadas: estudo de caso em experimento com brócolis. *Bragantia* 69: 1035-1046.
- LÚCIO AD; COUTO MRM; LOPES SJ; STORCK L. 2011. Transformação box-cox em experimentos com pimentão em ambiente protegido. *Horticultura Brasileira* 29: 38-42.
- LÚCIO AD; SOUZA MF; HELDWEIN AB; LIEBERKNECHT D; CARPES RH; CARVALHO MP. 2003. Tamanho da amostra e método de amostragem para avaliação de características do pimentão em estufa plástica. *Horticultura Brasileira* 21: 180-184.
- MARQUES DG; STORCK L; LOPES SJ; MARTIN TN. 2000. Qualidade dos ensaios de competição de cultivares de milho no estado do rio grande do sul. *Ciência Rural* 30: 381-385.
- MARTIN TN; STORCK L. 2008. Análise das pressuposições do modelo matemático em experimentos agrícolas no delineamento blocos ao acaso. In: MARTIN TN & ZIECH MF (org). *Sistemas de Produção Agropecuária*. Curitiba: UTFPR. p. 177-196.
- MORENO JA. 1961. *Clima no Rio Grande do Sul*. Porto Alegre: Secretaria da Agricultura. 41p.
- SISTEMA Brasileiro de Classificação de Solos. Rio de Janeiro: Embrapa Solos, 2006. 306p.
- SOCIEDADE BRASILEIRA DE CIÊNCIA DO SOLO. Comissão de química e fertilidade do solo. 2004. *Manual de adubação e de calagem para os estados do Rio Grande do Sul e de Santa Catarina*. Porto Alegre: Núcleo Regional Sul - Sociedade Brasileira de Ciência

- do Solo. 394p.
- SPRENT P; SMEETON NC. 2007. *Applied nonparametric statistical methods*. Boca Raton: Chapman & Hall. 530p.
- STEEL RGD; TORRIE JH; DICKEY DA. 1997. *Principles and procedures of statistics: a biometrical approach*. New York: McGraw-Hill. 666p.
- STORCK L; GARCIA DC; LOPES SJ; ESTEFANEL V. 2011. *Experimentação vegetal*. Santa Maria: UFSM. 200p.
- STORCK L; LOPES SJ; MARQUES DG; TISOTT CA; ROS CA. 2000. Análise de covariância para melhoria da capacidade de discriminação em ensaios de cultivares de milho. *Pesquisa Agropecuária Brasileira* 35: 1311-1326.
- ZANARDO B; LÚCIO AD; PALUDO AL; SANTOS D; BENZ V. 2010. Posições das mudas de alface nas bandejas de poliestireno e efeitos na normalidade e homogeneidade dos erros na produção de plantas. *Revista Ciência Agronômica* 41: 285-293.
-