

UNIVERSIDADE FEDERAL DO PARANÁ
SETOR DE CIÊNCIAS AGRÁRIAS

**MÉTODOS ESTATÍSTICOS PARA ANÁLISE DE DADOS DE CONTAGEM DE
ÁCAROS RAJADOS E TEORES FOLIARES DE NUTRIENTES EM POMARES DE
PESSEGUEIRO**

CURITIBA

2008

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

NERIO APARECIDO CARDOSO

**MÉTODOS ESTATÍSTICOS PARA ANÁLISE DE DADOS DE CONTAGEM DE
ÁCAROS RAJADOS E TEORES FOLIARES DE NUTRIENTES EM POMARES DE
PESSEGUEIRO**

Dissertação apresentada ao Programa de Pós-Graduação em Agronomia, área de concentração em Produção Vegetal, Departamento de Fitotecnia e Fitossanitarismo, Setor de Ciências Agrárias, Universidade Federal do Paraná, como parte das exigências para a obtenção do título de Mestre em Ciências.

Orientador: Dr. Lino Bittencourt Monteiro
Co-Orientador: PhD. Antônio Carlos Vargas Motta
Co-Orientador: Dr. Joel Maurício Corrêa da Rosa
Co-Orientadora: Dra. Louise Larissa May de Mio

CURITIBA

2008

DEDICATÓRIA

À minha mãe:

*Onilda da Aparecida Cardoso, por existir e pelas sabias palavras de conforto nos momentos
mais difíceis de minha existência.*

e

À minha família:

*Inês de Fátima da Silva e Bruno da Silva Cardoso, pessoas que me deram muito amor
e carinho para que eu pudesse vencer mais este desafio!*

AGRADECIMENTOS

A DEUS, por ter me guiado pelos caminhos da vida!

Ao Professor Dr. Lino Bittencourt Monteiro da Universidade Federal do Paraná pela amizade, companheirismo e orientação.

Ao Professor Dr. Antônio Carlos Vargas Motta, Coordenador do Departamento de Fitotecnia e Fitossanitarismo da UFPR, pela co-orientação, amizade e o apoio nos momentos difíceis durante minha caminhada no setor Agrário.

À Professora Dra. Beatriz Monte Serrat Prevedello, do Departamento de Fitotecnia e Fitossanitarismo da UFPR, pela co-orientação e apoio nos momentos difíceis, não tenho dúvidas da contribuição significativa e incondicional neste trabalho, uma amável pessoa com uma visão incrível.

Ao Professor Dr. Joel Maurício Corrêa da Rosa, Coordenador do Departamento de Estatística da UFPR, pela co-orientação e principalmente pela amizade cultivada desde a minha formação na graduação.

A Professora Dra. Francine Cuquel, Coordenadora do Departamento de Pós-Graduação em Agronomia – Produção Vegetal da UFPR, pelas palavras amigas nos momentos de dificuldades, tem infinito carinho por esta pessoa maravilhosa.

Ao Professor Dr. Anselmo Chaves Neto, Coordenador do Departamento de Pós-Graduação em Métodos Numéricos em Engenharia da UFPR, pelas palavras de apoio que me encorajaram a realizar este trabalho e principalmente pela sua amizade compartilhada com muitos cafezinhos da cantina.

Ao colega e amigo Professor Dr. Renato Tratch, Diretor Acadêmico e de Pesquisa da Fazenda Experimental Gralha Azul – Pontifícia Universidade Católica do Paraná, pelo fornecimento dos dados que gerou este trabalho e pelos sábios conselhos dados e principalmente pela amizade e finalmente pelos diversos lanches realizados em sua casa, quando estávamos trabalhando.

Esta pessoa é muito especial, a Professora Dra. Nilce Nazareno da Fonte, Coordenadora do Laboratório de Farmacognosia – Departamento de Farmácia da UFPR, pelas palavras amigas nos momentos complexos de minha formação, mudando completamente minha maneira de ver o mundo, ou seja, me deu outros óculos para ver a vida, tenho infinito carinho por esta pessoa maravilhosa.

A Professora Dra. Louise Larissa May de Mio do Departamento de Fitotecnia e Fitossanitarismo da UFPR, pela co-orientação e apoio nos momentos difíceis durante minha caminhada no setor Agrário.

Ao Msc. Alex Sandro Poltronieri e Msc. Josélia Schuber, pelas valiosas contribuições para aprimoramento desse trabalho e certo de que nessas pessoas sempre encontrarei a amizade verdadeira.

Aos estagiários do Laboratório de Manejo Integrado de Pragas (UFPR): Emily, Eder, Daiane, Giovana, Deise, Taciana, Maísa e aos funcionários, Cecília, Neli e Machado pelo incentivo e amizade.

Ao Mestrando Eng^o Agrônomo Eliseu, pela amizade e pelos momentos que confabulamos.

Ao Doutorando e Msc. Eng^o Agrônomo, Patrik Luis Pastori, pela amizade e pelas sabias palavras escritas e pronunciadas.

À Universidade Federal do Paraná e ao Programa de Pós-Graduação em Produção Vegetal pela oportunidade concedida.

Aos Professores do Programa de Pós-Graduação em Produção Vegetal.

E a todas as pessoas que de alguma maneira contribuíram para a realização desse trabalho.

“O grau de dificuldades que passamos em nossas vidas, depende do ponto de vista que encaramos nossos problemas”

Nerio Ap. Cardoso

SUMÁRIO

LISTA DE FIGURAS	IX
LISTA DE TABELA.....	XII
LISTA DE QUADROS	XV
INTRODUÇÃO GERAL	1
REFERÊNCIAS	4
CAPÍTULO I - FUNDAMENTAÇÃO TEÓRICA DOS MÉTODOS ESTATÍSTICOS	6
1 INTRODUÇÃO.....	9
2 DEFINIÇÃO DE VARIÁVEL.....	10
2.1 VARIÁVEL ALEATÓRIA DISCRETA	10
2.2 VARIÁVEL ALEATÓRIA CONTÍNUA.....	10
3 MENSURAÇÃO DE VARIÁVEIS	11
3.1 ESCALA NOMINAL.....	11
3.2 ESCALA ORDINAL.....	11
3.3 ESCALA DE RAZÃO.....	11
4 MEDIDAS DE TENDÊNCIA CENTRAL	11
4.1 MÉDIA ARITMÉTICA	12
4.2 MEDIANA	12
4.3 MODA.....	12
5 MEDIDA DE DISPERSÃO	12
5.1 AMPLITUDE TOTAL	13
5.2 DESVIO-PADRÃO.....	13
5.3 COEFICIENTE DE VARIAÇÃO.....	13
6 APRESENTAÇÃO GRÁFICA	14
7 ESTATÍSTICA NÃO PARAMÉTRICA	14
8 TESTES NÃO PARAMÉTRICOS	15
9 TESTE DE HIPÓTESE.....	16
10 MEDIDAS DE CORRELAÇÃO.	16
10.1 ANÁLISE DE CONTINGÊNCIA C	17
10.2 ANÁLISE DE CORRELAÇÃO DE SPEARMAN.....	17
10.3 ANÁLISE DE CORRELAÇÃO DE PEARSON.....	18
10.4 ANÁLISE DE CORRELAÇÃO CANÔNICA	19
11 ANÁLISE DE VARIÂNCIA	22

11.1	PARCELAS SUB-SUB-DIVIDIDAS.....	22
11.2	MODELO MATEMÁTICO.....	23
11.3	DECOMPOSIÇÃO DAS CAUSAS DE VARIAÇÃO.....	24
11.4	SOMA DE QUADRADOS.....	25
11.5	TESTE DE COMPARAÇÃO DE MÉDIAS.....	26
11.6	TRANSFORMAÇÃO BOX-COX.....	27
12	MODELOS LINEARES GENERALIZADOS.....	27
12.1	DEFINIÇÃO.....	28
12.2	LIGAÇÕES CANÔNICAS.....	30
12.3	ESTRUTURA DE CORRELAÇÃO.....	31
12.4	TESTES DE HIPÓTESES.....	32
13	ANÁLISE DE CLASSIFICAÇÃO.....	32
13.1	REDES NEURAS.....	33
13.2	ANÁLISE DISCRIMINANTE.....	33
13.3	ANÁLISE DE REGRESSÃO.....	33
13.4	REGRESSÃO LOGÍSTICA.....	34
13.5	ANÁLISE DE CLUSTER.....	34
13.6	ANÁLISE DE COMPONENTES PRINCIPAIS.....	34
13.7	ÁRVORES DE CLASSIFICAÇÃO.....	35
13.7.1	INTERPRETAÇÃO.....	35
13.7.2	VANTAGENS.....	35
13.7.3	DESVANTAGENS.....	36
13.7.4	DEFINIÇÃO.....	36
14	BOOTSTRAPING.....	37
15	SOFTWARE WEKA.....	38
15.1	MÉTODOS DE CLASSIFICAÇÃO.....	39
15.2	MÉTODOS DE PREDIÇÃO.....	39
15.3	INICIALIZAÇÃO DAS ANÁLISES.....	39
CAPÍTULO II - MODELOS LINEARES GENERALIZADOS, ÁRVORE DE CLASSIFICAÇÃO E ANOVA PARA ANÁLISE DE DADOS DE CONTAGEM DE ÁCARO RAJADO EM POMARES DE PESSEGUEIRO NA LAPA E CAMPO LARGO, PR.....		47
1	INTRODUÇÃO.....	50

2	MATERIAL E MÉTODOS.....	52
2.1	<i>MATERIAL</i>	52
2.2	<i>MÉTODOS</i>	55
3	RESULTADOS E DISCUSSÃO	57
3.1	<i>ANÁLISE DESCRITIVA</i>	57
3.2	<i>ANÁLISE DE DISPERSÃO</i>	59
3.3	<i>ANÁLISE DE VARIÂNCIA</i>	62
3.4	<i>ANÁLISE DOS PRESSUPOSTOS DA ANOVA</i>	63
3.5	<i>MODELOS LINEARES GENERALIZADOS PARA DADOS LONGITUDINAIS</i>	68
3.6	<i>ÁRVORE DE CLASSIFICAÇÃO</i>	70
4	CONCLUSÃO.....	72
CAPÍTULO III - TEORES FOLIARES EM POMARES DE PESSEGUEIRO SOB		
ADUBAÇÃO NITROGENADA E POTÁSSICA NA LAPA E CAMPO LARGO, PR.....		
1	INTRODUÇÃO.....	83
2	MATERIAL E MÉTODOS.....	86
2.1	<i>MATERIAL</i>	86
2.2	<i>MÉTODOS</i>	89
3	RESULTADOS E DISCUSSÃO	90
3.1	<i>ANÁLISE DE DISPERSÃO</i>	90
3.2	<i>COEFICIENTES DE VARIAÇÃO</i>	96
3.3	<i>ANÁLISE DE DADOS DISCREPANTES</i>	99
3.4	<i>ANÁLISE DE CORRELAÇÃO</i>	106
3.5	<i>ANÁLISE DE CORRELAÇÃO CANÔNICA</i>	109
3.6	<i>ANÁLISE DE VARIÂNCIA</i>	112
3.7	<i>ANÁLISE DE COMPARAÇÕES MÚLTIPLAS</i>	114
3.8	<i>ANÁLISE DE MODELOS LINEARES GENERALIZADOS</i>	121
4	CONCLUSÕES.....	127

LISTA DE FIGURAS

CAPÍTULO I - FUNDAMENTAÇÃO TEÓRICA DOS MÉTODOS ESTATÍSTICOS

FIGURA 1 - PASSOS PARA APLICAÇÃO DA TÉCNICA *BOOTSTRAP*.....37

CAPÍTULO II - MODELOS LINEARES GENERALIZADOS, ÁRVORE DE CLASSIFICAÇÃO E ANOVA PARA ANÁLISE DE DADOS DE CONTAGEM DE ÁCARO RAJADO EM POMARES DE PESSEGUEIRO NA LAPA E CAMPO LARGO, PR.

FIGURA 1 - DISPOSIÇÃO DE PÊSSEGO EM SUBPARCELAS E SUB-SUBPARCELAS CONSTITUÍDA DE DOSES DE NITROGÊNIO E POTÁSSIO EM DOIS POMARES, SITUADOS NA LAPA E CAMPO LARGO (PR), 2003 a 2006. .54

FIGURA 2 - BOX-PLOT DO NÍVEL POPULACIONAL DE ÁCAROS MÓVEIS SOB AS DIFERENTES DOSES DE ADUBAÇÕES NITROGENADAS E POTÁSSICAS ($\text{kg}\cdot\text{ha}^{-1}\cdot\text{ano}^{-1}$) REALIZADAS NAS SAFRAS DE 2003/04, 2004/05 E 2005/06 NO POMAR DE PÊSSEGO. LAPA, PR. 61

FIGURA 3 - ANÁLISE DOS PRESSUPOSTOS DA ANOVA SEM TRANSFORMAÇÃO DOS DADOS, LAPA, PR..... 64

FIGURA 4 - ANÁLISE DOS PRESSUSPOSTOS DA LAPA, COM TRANSFORMAÇÃO BOX-COX $\lambda=0,02020202$ 65

FIGURA 5 - ANÁLISE DOS PRESSUPOSTOS DA ANOVA SEM TRANSFORMAÇÃO, CAMPO LARGO..... 66

FIGURA 6 - ANÁLISE DOS PRESSUSPOSTOS DA ANOVA COM TRANSFORMAÇÃO BOX-COX $\lambda= 0,06060606$, CAMPO LARGO..... 67

FIGURA 7 - ÁRVORE DE CLASSIFICAÇÃO COM 91,48% DE CONFIABILIDADE ORIGINADA DO TEOR FOLIAR E ADUBAÇÕES DE NITROGENIO E POTASSIO EM RELAÇÃO AO NÍVEL DE ÁCAROS RAJADOS AMOSTRADO NO POMAR DE PESSEGUEIRO LOCALIZADO NA LAPA, PR..... 70

CAPÍTULO III - TEORES FOLIARES EM POMARES DE PESSEGUEIRO SOB ADUBAÇÃO NITROGENADA E POTÁSSICA NA LAPA E CAMPO LARGO

- FIGURA 1 - DISPOSIÇÃO DE PÊSSEGO EM SUBPARCELAS E SUB-SUBPARCELAS CONSTITUÍDA DE DOSES DE NITROGÊNIO E POTÁSSIO EM DOIS POMARES, SITUADOS NA LAPA E CAMPO LARGO, PR, 2003 a 2006.... 88
- FIGURA 2 - DISPERSÃO DE MACRONUTRIENTES, OBSERVADOS EM DIAGONE FOLIAR, NA PRESENÇA DE ADUBAÇÃO NITROGENADA ($\text{kg}\cdot\text{ha}^{-1}\text{ ano}^{-1}$), APLICADA EM POMARES DE PÊSSEGO 'CHIMARRITA'. LAPA, PR. NAS SAFRAS 2003/04, 2004/05 E 2005/06..... 91
- FIGURA 3 - DISPERSÃO DE MICRONUTRIENTES, OBSERVADOS NOS TEORES FOLIARES, NA PRESENÇA DE ADUBAÇÃO NITROGENADA ($\text{kg}\cdot\text{ha}^{-1}\text{ ano}^{-1}$), APLICADA EM POMARES DE PÊSSEGO 'CHIMARRITA'. LAPA, PR. SAFRAS 2003/04, 2004/05 E 2005/06..... 92
- FIGURA 4 - DISPERSÃO DE MACRONUTRIENTES, OBSERVADOS NOS TEORES FOLIARES, NA PRESENÇA DE ADUBAÇÃO NITROGENADA ($\text{kg}\cdot\text{ha}^{-1}\text{ ano}^{-1}$), APLICADA EM POMARES DE PÊSSEGO 'CHIMARRITA'. CAMPO LARGO, PR. SAFRAS 2003/04 E 2004/05. 94
- FIGURA 5 - DISPERSÃO DE MICRONUTRIENTES, OBSERVADOS NOS TEORES FOLIARES, NA PRESENÇA DE ADUBAÇÃO NITROGENADA ($\text{kg}\cdot\text{ha}^{-1}\text{ ano}^{-1}$), APLICADA EM POMARES DE PÊSSEGO 'CHIMARRITA'. CAMPO LARGO, PR. SAFRAS 2003/04 E 2004/05. 95
- FIGURA 6 - *BOX-PLOT* DOS MACRONUTRIENTES, NOS DIFERENTES DOSES DE ADUBAÇÕES NITROGENADAS ($\text{kg ha}^{-1}\text{ ano}^{-1}$) REALIZADAS NO POMAR DE PESSEGUEIROS 'CHIMARRITA'. LAPA, PR. 100
- FIGURA 7 - *BOX-PLOT* DOS MICRONUTRIENTES, NOS DIFERENTES DOSES DE ADUBAÇÕES NITROGENADAS ($\text{kg ha}^{-1}\text{ ano}^{-1}$) REALIZADAS EM POMAR DE PÊSSEGO 'CHIMARRITA'. LAPA, PR..... 101

FIGURA 8 - <i>BOX-PLOT</i> DOS MACRONUTRIENTES, NOS DIFERENTES DOSES DE ADUBAÇÕES NITROGENADAS ($\text{kg ha}^{-1} \text{ano}^{-1}$) REALIZADAS NO POMAR DE PÊSSEGO 'CHIMARRITA' EM CAMPO LARGO, PR.	103
FIGURA 9 - BOX PLOT DOS MICRONUTRIENTES, NOS DIFERENTES DOSES DE ADUBAÇÕES NITROGENADAS ($\text{kg ha}^{-1} \text{ano}^{-1}$) REALIZADAS NO POMAR DE PÊSSEGO 'CHIMARRITA' EM CAMPO LARGO, PR.	104
FIGURA 10 - ANÁLISE DE CORRELAÇÃO DE NPK, SAFRAS 2003/04, 2004/05 E 2005/06, NA PRESENÇA DE DIFERENTES DOSES DE ADUBAÇÕES NITROGENADAS E POTÁSSICA ($\text{kg ha}^{-1} \text{ano}^{-1}$) DOS POMARES DE PÊSSEGO 'CHIMARRITA'. LAPA, PR.	107
FIGURA 11 - ANÁLISE DE CORRELAÇÃO DE NPK, SAFRAS 2003/04 E 2004/05, NA PRESENÇA DE DIFERENTES DOSES DE ADUBAÇÕES NITROGENADAS E POTÁSSICA ($\text{kg ha}^{-1} \text{ano}^{-1}$) DOS POMARES DE PESSEGUEIRO 'CHIMARRITA'. CAMPO LARGO, PR.	108
FIGURA 12 - ANÁLISE DE CORRELAÇÃO CANÔNICA DOS TEORES FOLIARES DE MACRO E MICRO NUTRIENTES DAS SAFRAS DE 2003/04, 2004/05, 2005/06 EM POMAR DE PÊSSEGO 'CHIMARRITA'. LAPA, PR.	110
FIGURA 13 - ANÁLISE DE CORRELAÇÃO CANÔNICA DA DIAGNOSE FOLIAR DAS SAFRAS DE 2003/04 E 2004/05 EM POMAR DE PÊSSEGO 'CHIMARRITA'. CAMPO LARGO, PR.	111
FIGURA 14 – GRÁFICOS DE INTERAÇÃO DA CONCENTRAÇÃO DE NITROGÊNIO NA FOLHA ENTRE ADUBAÇÃO NITROGENADA, POTÁSSICA E SAFRAS NO POMAR DE PÊSSEGO 'CHIMARRITA', LOCALIZADO EM CAMPO LARGO, PR.	118
FIGURA 15 – GRÁFICOS DE INTERAÇÃO DA CONCENTRAÇÃO DE FÓSFORO NA FOLHA ENTRE ADUBAÇÃO NITROGENADA E POTÁSSICA NO POMAR DE PÊSSEGO 'CHIMARRITA', LOCALIZADO EM CAMPO LARGO, PR.	119

LISTA DE TABELA

CAPÍTULO II - MODELOS LINEARES GENERALIZADOS, ÁRVORE DE CLASSIFICAÇÃO E ANOVA PARA ANÁLISE DE DADOS DE CONTAGEM DE ÁCARO RAJADO EM POMARES DE PESSEGUEIRO NA LAPA E CAMPO LARGO, PR.

- TABELA 1 - ANÁLISE QUÍMICA DO SOLO EM POMAR DE PESSEGUEIRO NO PERIMETRO DO TRONCO (PT), NA PROJEÇÃO DA COPA (PC), A 0 - 20 CM DE PROFUNDIDADE, EM 2004 (LAPA - PR).53
- TABELA 2 - ANÁLISE QUÍMICA DO SOLO EM POMAR DE PESSEGUEIRO NO PERIMETRO DO TRONCO (PT), NA PROJEÇÃO DA COPA (PC) E NA ENTRE LINHA (EL), A 0 - 20 E 20 - 40 CM DE PROFUNDIDADE, EM 2003 (CAMPO LARGO - PR). 53
- TABELA 3 - DESCRITIVO DO NÍVEL POPULACIONAL DE ÁCAROS MÓVEIS SOB AS DIFERENTES DOSES DE ADUBAÇÕES NITROGENADAS E POTÁSSICAS ($\text{kg}\cdot\text{ha}^{-1}\text{ano}^{-1}$), E BLOCOS, REALIZADAS NAS SAFRAS DE 2003/04, 2004/05 E 2005/06 NO POMAR DE PÊSSEGO. LAPA E CAMPO LARGO, PR.58
- TABELA 4 - ANÁLISE DE VARIÂNCIA (ANOVA) DO NÍVEL POPULACIONAL DE ÁCAROS MÓVEIS SOB AS DIFERENTES DOSES DE ADUBAÇÕES NITROGENADAS E POTÁSSICAS ($\text{kg ha}^{-1}\text{ano}^{-1}$), E BLOCOS, REALIZADAS NAS SAFRAS DE 2003/04, 2004/05 E 2005/06 NO POMAR DE PÊSSEGO DA LAPA E CAMPO LARGO, PR. 62
- TABELA 5 - ESTIMATIVA PARA O NÍVEL POPULACIONAL DE ÁCAROS MÓVEIS SOB AS DIFERENTES DOSES DE ADUBAÇÕES NITROGENADAS E POTÁSSICAS ($\text{K g ha}^{-1}\text{ano}^{-1}$) REALIZADAS NAS SAFRAS DE 2003/04, 2004/05 E 2005/06 EM POMARES DE PÊSSEGO DA LAPA E CAMPO LARGO, PR. 69

CAPÍTULO III - TEORES FOLIARES EM POMARES DE PESSEGUEIRO SOB ADUBAÇÃO NITROGENADA E POTÁSSICA NA LAPA E CAMPO LARGO

- TABELA 1 - ANÁLISE QUÍMICA DO SOLO EM POMAR DE PESSEGUEIRO NO PERIMETRO DO TRONCO (PT), NA PROJEÇÃO DA COPA (PC), A 0 - 20 CM DE PROFUNDIDADE, EM 2004 (LAPA - PR)..... 86
- TABELA 2 - ANÁLISE QUÍMICA DO SOLO EM POMAR DE PESSEGUEIRO NO PERIMETRO DO TRONCO (PT), NA PROJEÇÃO DA COPA (PC) E NA ENTRE LINHA (EL), A 0 - 20 E 20 - 40 CM DE PROFUNDIDADE, EM 2003 (CAMPO LARGO - PR). 87
- TABELA 3 – COEFICIENTE DE VARIAÇÃO DOS MACRONUTRIENTES, NOS DIFERENTES DOSES DE ADUBAÇÕES NITROGENADAS REALIZADAS NAS SAFRAS DE 2003/04, 2004/05 e 2005/06 EM POMARES DE PÊSSEGO ‘CHIMARRITA’. LAPA E CAMPO LARGO, PR. 97
- TABELA 4 - COEFICIENTES DE VARIAÇÃO (CV) DOS MICRONUTRIENTES, NOS DIFERENTES DOSES DE ADUBAÇÕES NITROGENADAS REALIZADAS NAS SAFRAS 2003/04, 2004/05 e 2005/06 EM POMAR DE PÊSSEGO ‘CHIMARRITA’ LAPA E CAMPO LARGO, PR. 98
- TABELA 5 - ANÁLISE DE VARIÂNCIA DOS RESULTADOS DE NITROGENIO ($N \text{ g.kg}^{-1}$), FÓSFORO ($P \text{ g.kg}^{-1}$) E POTÁSSIO ($K \text{ g.kg}^{-1}$) ENCONTRADO NAS FOLHAS, EM PRESENÇA DE DIFERENTES DOSES DE ADUBAÇÃO NITROGENADA E POTÁSSICA, NO POMAR DE PÊSSEGO ‘CHIMARRITA’, LOCALIZADO NA LAPA E CAMPO LARGO, PR. 113
- TABELA 6 - ANÁLISE DOS RESULTADOS DE NITROGENIO ($N \text{ g.kg}^{-1}$), FÓSFORO ($P \text{ g.kg}^{-1}$) E POTÁSSIO ($K \text{ g.kg}^{-1}$) ENCONTRADOS NA FOLHA, EM PRESENÇA DE DIFERENTES DOSES DE ADUBAÇÃO NITROGENADA E POTÁSSICA, NO POMAR DE PÊSSEGO ‘CHIMARRITA’, LOCALIZADO NA LAPA, PR..... 116

TABELA 7 - ANÁLISE DOS RESULTADOS DE NITROGENIO ($N \text{ g.kg}^{-1}$), FÓSFORO ($P \text{ g.kg}^{-1}$) E POTÁSSIO ($K \text{ g.kg}^{-1}$) ENCONTRADOS NA FOLHA, EM PRESENÇA DE DIFERENTES DOSES DE ADUBAÇÃO NITROGENADA E POTÁSSICA, NO POMAR DE PÊSSEGO 'CHIMARRITA', LOCALIZADO 57442.146571(-)-157426(999(A)36.)1.57564(M)-1.52434(A)1.57564(R)-B.

LISTA DE QUADROS

CAPÍTULO I - FUNDAMENTAÇÃO TEÓRICA DOS MÉTODOS ESTATÍSTICOS

QUADRO 1 – DESCRIÇÃO DA ANÁLISE DE VARIÂNCIA PARA O DELINEAMENTO DE PARCELA SUBSUBDIVIDIDA.....	24
---	----

INTRODUÇÃO GERAL

Os dados obtidos em pesquisas agrícolas e biológicas entre outras áreas de pesquisas como ciências exatas e humanas podem ser submetidos a diversos métodos de análises estatísticas. Entretanto para submeter os dados aos métodos de análises, requer o conhecimento prévio da natureza dos dados (SANTOS, MOREIRA E BELTRÃO, 1998), para delimitar os métodos estatísticos aplicáveis ao conjunto de dados.

Segundo CAMPOS et. al. (2007), é importante a identificação das características e tipo da variável envolvida no experimento, essas informações delimitam os métodos estatísticos mais adequados, e à interpretação correta dos resultados da pesquisa. Portanto, para uma aumentar a precisão das análises estatísticas, bem como realizar conclusões coerentes é necessário que o analista conheça a natureza e as relações entre as variáveis.

Outro ponto a ser verificado é o relacionamento entre as variáveis que freqüentemente é inevitável em virtude da natureza única das variáveis, isto leva a simplificação, muitas vezes, grosseira com a finalidade de facilitar a análise de dados ou torná-la viável (NETO, 2007). Para analogia entre as variáveis de forma sistêmica, procura-se a relação causa e efeito entre as variáveis (MATTAR, 2001). Em geral, as diferenças existentes entre grupos ou populações de interesse, não estão sendo explicadas apenas por uma variável e sim por conjunto delas, podendo esta relação ser explicada através de modelos matemáticos (FAYYAD, 1996).

O estudo de uma única variável indica que determinado tratamento é mais adequado, porém ao considerar outras variáveis relacionadas à variável de interesse, outros tratamentos podem ser mais indicados. Isso se deve pela interação entre as variáveis (PONTES, 2005).

Dentre diversos métodos estatísticos, o mais utilizado em análises agronômicas é Análise de Variância (ANOVA) (DRAPER e SMITH, 1998). A ANOVA consiste em decompor a variação total das observações do experimento em partes que podem ser atribuídas a causas controladas (conhecidas) e em partes a causas não controladas e/ou não controláveis (desconhecidas), o erro ou resíduo. O delineamento abordado neste trabalho é em parcelas sub-sub-dividida que consiste em alocar o fator A em parcelas principais (ou primária), o fator B nas sub-parcelas (secundárias) e o fator C em sub-sub-parcelas. Esta aplicação decorre da facilidade de instalação do experimento (MONTGOMERY, 2001).

O erro ou resíduo pode ocorrer em função do material que se está trabalhando ou em função do ambiente em que o experimento é conduzido (MONTGOMERY, 2001). Uma fonte

de erro pode ser a maneira como o experimento é conduzido pelo experimentador, bem como variáveis que estão fora dos padrões ou são exceções às regras (ANJOS, 2004; FELDENS, 1997).

Segundo MONTGOMERY (2001), para aplicação da ANOVA é necessário a verificação dos pressupostos de normalidade, homocedasticidade, independência e aditividade dos resíduos.

Uma das formas de evitar vieses nas conclusões é planejar um experimento de forma a evitarem-se influências sistemáticas de variáveis sobre os tratamentos, invalidando os métodos de hipóteses. O uso de delineamentos experimentais permite que se utilizem simultaneamente vários tratamentos num mesmo experimento, dependendo das características das variáveis, sem invalidar as pressuposições exigidas pelo modelo matemático na ANOVA (ANJOS, 2004).

É comum a utilização da ANOVA em trabalhos científicos cuja mensuração variável resposta é por meio de contagem, por exemplo, monitoramento de ácaros por um determinado tempo é realizada por meio de amostragem aleatória em plantas e folhas cujas variáveis respostas são contagens. Portanto, é possível que a variável resposta, observada ao longo do tempo, tenha uma estrutura de correlação entre as datas em que se realizaram as amostragens. Isso tem que ser levado em consideração na estimação dos parâmetros (COSTA, 2003).

Quando se aplica o método ANOVA é desconsiderada a estrutura de correlação entre as observações, mensuradas ao longo do tempo, assumindo, também, que a distribuição da variável resposta é normal. Por outro lado, na análise de dados oriundos de contagem, proporções, é possível utilizar a teoria de modelos lineares generalizados (MCCULLAGH e NELDER, 1989; DEMETRIO, 2001).

Uma forma de analisar os dados é dispensar o conhecimento da forma e parâmetros das funções de distribuição da variável aleatória, possibilitando a aplicação de métodos não-paramétricos.

As vantagens na aplicação dos métodos não-paramétricos está no fato de não haver necessidade dos dados corroborarem com distribuições probabilísticas, variâncias podem ser heterogêneas (PONTES, 2005). Portanto não há necessidade dos dados atenderem os pressupostos exigidos pelos métodos de estimação de modelo matemático. Outro porém, os métodos não paramétricos pode ser aplicados a dados que não sejam numéricos (GOMES, 1990).

Dentre vários métodos não-paramétricos eficientes, o que se aplicou neste trabalho foi a Árvore de Classificação, o qual contém nós com decisões preditivas, auxiliando nas tomadas de decisões, por meio da lógica condicional.

A Árvore de Classificação seleciona o atributo mais significativo aos dados, sendo um modelo preditivo representado por um gráfico em forma de árvore contendo as decisões a serem inferidas indutivamente (FAYYAD, 1996). Esta forma de árvore tem um conjunto de nós de decisão que corresponde uma variável relevante, a qual pode subsidiar decisões depois de treinadas de acordo com um conjunto de treinos realizados na variável de interesse (CARVALHO, 2005). A decisão tomada é representada pela folha da árvore, representando valor predito da variável em estudo.

A árvore pode ser analisada a partir do topo do gráfico também chamado nó raiz, aonde se localiza o fator mais expressivo (ALVES e FERREIRA, 2007). As vantagens ao analisar conjunto de dados por meio de Árvores de Classificação estão em: robustez nos cálculos complexos; dados discrepantes não influenciam nas decisões, a falta de valores para alguns atributos não influencia nos valores preditivos; capacidade de aprender por meio de exemplos; facilidade na preparação dos dados para análises (GOEBEL e GRUENWALD, 1999). Entretanto, a desvantagem é o uso do algoritmo de aproximação que podem provocar desperdício de processamento (QUINLAN, 1998).

Para analisar dados através do método de Árvore de Classificação, uma das ferramentas de mineração de dados mais simples e largamente utilizadas é *software* Weka (*WAIKATO ENVIRONMENT FOR KNOWLEDGE ANALISYS*). O sistema Weka é um *software* livre (de código aberto) para mineração de dados, desenvolvido em Java disponível para Windows e Linux (WITTEN e FRANK, 2005).

Neste estudo, o conjunto de dados foi obtido pelo Grupo Ensino, Extensão e Pesquisa em Produção Integrada (GEEPPI), de dois experimentos realizados com a cultura do pessegueiro, submetida a diferentes doses de adubação nitrogenada e potássica. Foi realizado análises dos teores foliares e contagem de ácaros móveis. O objetivo geral deste trabalho é apresentar análises estatísticas adequadas para avaliar a relação entre as variáveis coletadas pelo grupo GEEPPI. Os objetivos específico é mostrar que ANOVA não é eficiente para analisar dados de contagem, apresentar análise descritiva para experimentos com variáveis discretas e contínuas, apresentar os diferentes métodos de análises de correlações e finalmente, comparar os resultados da ANOVA, Modelos Lineares Generalizados para Dados Longitudinais e Árvore de Classificação.

REFERÊNCIAS

ALVES, L.; FERREIRA, A. **Inteligência Artificial e Data Mining**. Salvador: Departamento de Ciências da Computação da Universidade Federal da Bahia (UFBA). Disponível em: <http://twiki.im.ufba.br/pub/MAT054/SemestreArtigos20052/Artigo.doc>. Acesso em 22 de maio de 2007.

ANJOS, A. **Notas de Aula da Disciplina de Planejamento de Experimentos II**. Curitiba: UFPR, 2004.

CAMPOS, G. M.; BARBIN, L. E.; SPANÓ, E. C. J.; SILVA, S. R.; PÉCORÁ, D. J. **Estatística Prática para Docentes e Pós-Graduandos: tipos de variáveis**. Disponível em: http://www.forp.usp.br/restauradora/gmc/gmc_livro/gmc_livro_cap02.html. Acesso em 23 de setembro de 2007.

CARVALHO, R. B. **Árvore de Decisão - Algoritmo Genético para tratar o problema de pequenos disjuntos em classificação de dados**. Rio de Janeiro, 2005, 173 f. Tese (Doutorado em Ciências em Engenharia Civil) - Universidade Federal do Rio de Janeiro, (UFRJ).

COSTA, N. P. O. **Estatística**. São Paulo: Edgard Blucher, 2003.

DEMETRIO, C. G. B. **Modelos lineares Generalizados em experimentação agronômica**. Piracicaba: ESALQ, 2001. 113p.

DRAPER, N. R.; SMITH JR., H. **Applied Regression Analysis**. 3.ed.. New York: John Wiley & Sons, 1998. 706 p.

FAYYAD, U. **Advances in Knowledge Discovery and Data Mining** - Mit Press 1 ed. 1996.

FELDENS, M. A. **Knowledge discovery in databases**. 1997. Disponível em <http://www.ufrgs.br/feldens>. Acesso em: 20 de Dezembro de 2007.

GOEBEL, M.; GRUENWALD, L. **A survey of data mining and knowledge discovery software tools**. SIGKDD Explorations, v. 1, 1999. p. 20-33.

GOMES, P. F. **Curso de estatística experimental**. 13 ed. Piracicaba: ESALQ-USP, 1990.

MATTAR, F. N. **Pesquisa de marketing: metodologia, planejamento, execução e análise**. 2. ed. São Paulo: Atlas, 1994. 275 p.

MCCULLAGH, P.; NELDER, J. A. **Generalized linear models**. 2.ed. London: Chapman & Hall, 1989. 511 p.

MONTGOMERY, D. C. **Design and analysis of experiments**. 3ed. New York: John Wiley & Sons, 2001. 856 p.

NETO, J. M. M. **Estatística multivariada**: uma visão didático-metodológica. Disponível em <http://criticanarede.com/cien_estatistica.html>. Acesso em 23 de fevereiro de 2007.

PONTES, A. C. F. **Análise de variância multivariada com a utilização de métodos não-paramétricos e componentes principais baseados em matrizes de postos**. Piracicaba, 2005, 117 f. Tese (Doutorado em Estatística e Experimentação Agronômica) - ESALQ, Universidade de São Paulo.

QUINLAN, R. C **4.5**: Programs for Machine Learning. San Francisco: Morgan Kaufmann, 1993.

SANTOS, W. J.; MOREIRA, N. A. J.; BELTRÃO, M. E. N. Avaliação do Emprego dos Testes de Comparação de Médias na Revista Pesquisa Agropecuária Brasileira (PAB) de 1980 a 1994. Brasília. **Revista Pesquisa Agropecuária Brasileira (PAB)**, v33, n3, p225-230, Mar.1998

WEKA MACHINE LEARNING PROJECT. Disponível em: <<http://www.cs.waikato.ac.nz/~ml>>. Acesso em: 10 de abril de 2007.

WEKA SOFTWARE. **Weka 3**: Data Mining with Open Source Machine Learning Software in Java. Disponível em: < <http://www.cs.waikato.ac.nz/~ml/weka/index.html> >Acesso em: 10 de abril de 2007.

WITTEN, I. H.; FRANK, E. **Data Mining**: Practical machine learning tools with Java implementations. Morgan Kaufmann: San Francisco, 2000.

CAPÍTULO I - FUNDAMENTAÇÃO TEÓRICA DOS MÉTODOS ESTATÍSTICOS

FUNDAMENTAÇÃO TEÓRICA DOS MÉTODOS ESTATÍSTICOS

RESUMO

Freqüentemente observa-se em experimentos agrícola há necessidade da utilização de métodos estatísticos para fundamentar conclusões agronômicas, entretanto para utilizar métodos estatísticos se faz necessário realizar um estudo prévio das variáveis, dos métodos estatísticos, bem como a aplicabilidade. Entretanto há dificuldade de realizar a fundamentação teórica devido à complexidade ou pelo desconhecimento da necessidade de verificação dos pressupostos dos métodos estatísticos. Portanto, é freqüente a utilização incorreta dos métodos estatísticos. Entretanto, é fundamental para se obter resultados de análises estatísticas fidedigna aos dados a fundamentação teórica dos métodos estatísticos empregados no estudo. Com isso, o objetivo deste capítulo é a fundamentação teórica das principais metodologias estatísticas empregadas para análises de contagens de ácaros e teores foliares sobre adubação nitrogenada e potássica abordados nos capítulos 2 e 3. Desta maneira, o capítulo I aborda a caracterização das variáveis aleatórias, a classificação dos dados quanto a sua mensuração, medidas de tendência central e dispersão e suas vantagens de desvantagens, forma de apresentação gráfica dos dados. Apresenta-se a definição de estatística não paramétrica e os principais testes bem como seus pressupostos para utilização da metodologia. Também são caracterizadas as hipóteses nulas e alternativas. Aborda as principais medidas de correlação e os pré-requisitos para aplicação dos métodos. Análise de Variância apresentada para delineamento em parcelas sub-subdividida. O Modelo Linear Generalizado apresentado neste estudo aborda sua definição, ligações canônicas, estruturas de correlação e testes de hipótese. São apresentados os principais métodos de classificação, realizando um estudo aprofundado em Árvore de Classificação. Também é realizado um estudo sobre *Bootstraping* e o *software* WEKA de distribuição gratuita.

Palavra-Chaves: Métodos estatísticos, Árvore de Classificação, Parcelas Sub-subdividida

RATIONALE THEORY OF STATISTICAL METHODS

ABSTRACT

In some cases, realizes that in agricultural experiments need to use statistical methods to support the conclusions agronomic, however to use statistical techniques is necessary to conduct a preliminary study on the variables, applicability of the assumptions and statistics tools. However there are difficulties in achieving the theoretical foundation due to the complexity or the ignorance of the need to verify the assumptions of the statistical methods. Therefore, it is often the wrong use of the statistical methods. Thus, objective of this chapter is the theoretical foundation of the main statistical methodologies employed for analysis of count for mites, leaf levels and nitrogen and potassium fertilization, discussed in Chapters 2 and 3. Thus the chapter 1 deals the characterization of random variables, the classification of data about its measurement, measures of central position and dispersion of its advantages and disadvantages, form of graphical presentation of data. It is the definition of non-parametric statistical tests and key and its assumptions for use of the methodology. It is also characterized the null hypotheses and alternative hypothesis. It addresses the main measures of correlation and the prerequisites for application of the methods. Analysis of Variance presented to design split-split-plot. Linear Model Generalized presented in this study addresses its definition, links canonical, structures of correlation and tests of hypothesis. Furthermore, the main methods of classification, conducting an in-depth study on the classification tree. It is also conducted a study on Bootstrapping and software WEKA free distribution.

Key-words: Split-split-plot, Classification Tree, Analysis of correlation.

1 INTRODUÇÃO

É constante a utilização de termos estatísticos nos mais diversos setores, como nas escolas, faculdades, empresas, entre outros. Mas é freqüente o desconhecimento do significado da palavra. No plural, estatísticas indicam qualquer coleção consistente de dados numéricos, reunidos com a finalidade de fornecer informações acerca de uma atividade qualquer, por exemplo, estatísticas demográficas (nascimento, falecimentos, matrimônios, etc.). No singular, indica atividade humana especializada ou um corpo de técnicas, metodologia desenvolvida para coletas de dados, classificação, a apresentação, análise e a interpretação de dados quantitativos e utilização desses dados para uma tomada de decisões (TOLEDO e OVALLE, 1995).

A estatística pode ser trabalhada na forma descritiva e/ou indutiva. A estatística descritiva tem como objetivo a observação de fenômenos, a organização e a classificação desses dados observados e a sua apresentação através de gráficos e tabelas, além do cálculo de coeficientes (estatísticas) que permitem descrever resumidamente os fenômenos (CAMPOS et. al., 2007). A estatística indutiva conhecida também como inferência estatística é um processo de generalização, a partir dos resultados particulares, ou seja, partindo da análise de uma amostra são obtidas conclusões representativas a população, entretanto, o processo de generalização está associado a incerteza, por estar trabalhando com uma parte (amostra) do total (população) de dados (SPIEGEL, 1994).

Nos trabalhos científicos, a necessidade de realizar conclusões para a população a partir de amostras mensuradas, por exemplo, em processos industriais, fenômenos físicos, biológicos ou sociais, que não se pode diretamente observar, é uma necessidade, em geral, constante (ALVES e CUNHA 2007). A pesquisa científica se constitui num processo iterativo de aprendizado. Para explicação de um fenômeno, coleta e analisa dados de acordo com uma hipótese. Por outro lado, a análise destes mesmos dados coletados por meio de amostragem ou experimentação geralmente sugere modificações da explicação do fenômeno (BOCK, 1975).

Outro, porém, com a velocidade com que se desenvolvem as inovações nas tecnologias de informática, cresce o número de *softwares*, *hardwares* entre outros equipamentos (CARVALHO, 2005), facilitando aplicação de diversas metodologias estatísticas. Entretanto, observa-se que as ferramentas estatísticas para análises dos dados são de uso corrente para a análise e interpretação de resultados, contudo, há situações em que não são utilizadas adequadamente. Um dos motivos é a diversidade de técnicas estatísticas

disponíveis atualmente (COSTA, 2003) implantadas em *softwares* estatísticos, sem a descrição dos pressupostos.

Assim, se torna necessário estudo teórico aprofundado da metodologia que se pretende utilizar nas análises de dados. Portanto, o objetivo deste capítulo é apresentar as principais metodologias estatísticas empregadas nos capítulos 2 e 3.

2 DEFINIÇÃO DE VARIÁVEL

O observador ao anotar ou medir a intensidade efetiva de um caráter variável em cada um dos objetos ou pessoas observadas obtêm uma variável. Por exemplo, ao registrar a idade das pessoas ao morrer, a estatura ou peso dos indivíduos, o rendimento das famílias em uma grande cidade, etc., estará obtendo um resultado (variável resposta). Estes resultados numéricos podem ser expressos de forma discreta ou descontínua e contínua (MATTAR, 2001).

2.1 VARIÁVEL ALEATÓRIA DISCRETA

Normalmente, a variável discreta resulta de contagem, razão pela qual seus valores são expressos através de números inteiros não-negativos. Pode variar em todo o conjunto de números inteiros racionais, e não apenas inteiros positivos de zero a infinito. Por exemplo, o interesse em saber o número de funcionários que praticam ginástica terapêutica em um determinado horário (COSTA, 1977).

2.2 VARIÁVEL ALEATÓRIA CONTÍNUA,

Formalmente, diz-se que X é uma variável contínua quando, ao passar de um valor real a para outro valor real b , pode assumir todos os valores intermediários entre a e b . Por exemplo, números de metros percorridos por um atleta durante certo período de tempo em uma pista circular.

Outra definição, geralmente utilizada em análises de decomposição da variação dos dados é variável independente e dependente ou resposta. Denomina-se por variável independente aquela cuja modificação se supõe poder produzir uma modificação num dado comportamento observável (variável dependente ou variável de resposta). Por exemplo, determinar a relação entre o número de erros na atividade laboral (variável dependente) e o número de treinamentos (variável independente) realizados nos funcionários; determinar o efeito do álcool sobre os reflexos motores (GOMES, 1990).

3 MENSURAÇÃO DE VARIÁVEIS.

A mensuração da variável resposta pode ser realizada em escalas nominal, ordinal e razão (BUSSAB e MORETTIN, 1987).

3.1 ESCALA NOMINAL

A escala nominal é a de mais baixo nível de mensuração. Utiliza símbolos, ou simplesmente números, para distinguir elementos em diferentes categorias. Geralmente não tem possibilidades de comparação de ordem crescente ou decrescente. Por exemplo: Masculino (M), Feminino (F); Perfeito (1), Defeituoso (0); Europeu (1), Americano (2), Africano (3), Asiático (4) (SIEGEL, 1975).

3.2 ESCALA ORDINAL

Escala ordinal utiliza números apenas para classificarmos elementos numa ordem crescente ou decrescente. Existe algum tipo de relação entre as categorias embora a diferença entre elas seja de difícil quantificação, exemplificando, classes sócio-econômicas (A, B, C, D, E); patentes do exército brasileiro (soldado, cabo, sargento, etc.); opinião sobre um determinado atendimento (Ruim, Regular, Bom, Muito bom, Excelente) (SIEGEL, 1975)

3.3 ESCALA DE RAZÃO

Equivalente a escala intervalar, tem-se a escala de razão, porém o valor zero é o verdadeiro ponto de origem, ponto de nulidade, ou seja, ponto conhecido. Importante a ser observado é o de que, nas escalas de razão, um valor de "2" efetivamente indica uma quantidade duas vezes maior do que a do valor "1", e assim por diante, por exemplo, a velocidade, tempo (SIEGEL, 1975).

4 MEDIDAS DE TENDÊNCIA CENTRAL

Medidas de tendência central são conhecidas como médias aritméticas, geométricas, harmônicas, quadráticas, entre outras, têm a função de descrever resumidamente uma distribuição de frequências. Portanto, apresentando o valor que melhor represente o conjunto de dados. Este conjunto de dados é formado por repetições da variável de interesse possibilitando verificar a dispersão dos dados indicando a qualidade da medida de tendência central.

4.1 MÉDIA ARITMÉTICA

A média aritmética de um conjunto de números é igual ao quociente entre a soma dos valores do conjunto e o número total de valores. Por exemplo:

$$\mu = \frac{\sum_{i=1}^n x_i}{N}$$

Esta medida é freqüentemente utilizada em diversos cálculos, e o cuidado que se deve ter com esta medida é verificar o quanto os valores extremos estão influenciando a média (ALBUQUERQUE, 2007).

4.2 MEDIANA

A mediana, também é uma medida de posição central, obtida pelo valor que separa os dados ordenados em duas partes iguais, metade dos dados tem valores maiores do que a mediana, a outra metade tem valores menores do que a mediana.

Há conjuntos de dados em que a média aritmética ou a mediana não é uma medida de tendência central que represente adequadamente os dados, o que pode ser explicado pela elevada variabilidade das variáveis ou alta freqüência de um determinado valor. Nestas condições sugere-se a moda (ALBUQUERQUE, 2007).

4.3 MODA

A moda é obtida pelo valor que ocorre com maior freqüência em um conjunto de dados.

Observa-se que as medidas centrais, mediana e moda, não levam em conta valores extremos, portanto seus parâmetros não são influenciados por valores discrepantes (ALBUQUERQUE, 2007).

5 MEDIDA DE DISPERSÃO

Um aspecto importante no estudo descritivo de um conjunto de dados é a verificação da variabilidade ou dispersão dos dados, em relação à medida de localização central da amostra (BARBIN, 2004). Variados métodos utilizados na Agronomia dependem da dispersão dos dados. Outro, porém, a representatividade da média associa-se a uma medida de dispersão. Algumas medidas são: amplitude total, desvio padrão e coeficiente de variação.

5.1 AMPLITUDE TOTAL

A medida de dispersão mais simples de ser obtida é amplitude total (MATTAR, 2001). Esta medida é indicada para dados com dispersão acentuada.

5.2 DESVIO-PADRÃO

Para obter uma medida da variabilidade ou dispersão na mesma escala de mensuração dos dados, tomamos a raiz quadrada da variância e obtemos o desvio padrão. As características contidas no desvio-padrão são as seguintes: pode assumir somente valores positivos, quanto maior mais dispersos os dados (MATTAR, 2001). Uma vantagem é que considera todos os valores, portanto, não perda de informação. Outro, porém, esta medida possibilita avaliar a qualidade do conjunto dos dados. É freqüente, em experimentos agrícolas, a comparação entre dois conjuntos de dados. Quando a dispersão dos dados é elevada, dificulta aplicação dos testes estatísticos paramétricos. Abaixo a função para calculo do desvio-padrão.

s = Desvio padrão amostral.

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

5.3 COEFICIENTE DE VARIAÇÃO

Quando o interesse é comparar a dispersão/desvios padrão de diferentes conjuntos, geralmente utiliza-se o coeficiente de variação (CV). O exemplo de uma aplicação seria o interesse de se conhecer a variação na altura das mulheres é maior ou menor do que as dos homens (BARBIN, 2004). A equação para calcular o CV é:

$$CV = \frac{s * 100}{\bar{x}}$$

Este método para comparar a dispersão dos dados é frequentemente utilizado em experimentos agrícolas. Entretanto é necessário verificar se as condições em que se obteve as variáveis de interesse é semelhante, para justificar uma possível diferença existente.

Entretanto é importante salientar que a dimensão dos CV's depende da natureza da variável resposta. Os CV's encontrados em variáveis geradas em ambientes controlados,

dependendo da natureza da variável, em geral é inferior aos encontrados em variáveis em ambientes não controlados. Da mesma forma ocorre com os interesses dos analistas de diferentes áreas de pesquisa.

6 APRESENTAÇÃO GRÁFICA

Estudos que baseiam suas conclusões em conjuntos de dados necessitam verificar a qualidade dos dados para validar os métodos que fundamentam as conclusões ou para encontrar problemas nos dados

É por meio de análises estatísticas exploratória, para verificar a qualidade dos dados, em que utilizam de desenhos esquemáticos (GONÇALVES e CASTRO, 1998), também conhecidos como diagramas, cartogramas, estereogramas e gráficos.

Os gráficos propriamente ditos são disponibilizados em diversos *softwares* em formato padrão, por exemplo, os gráficos setoriais, barras, linhas, *Box-plot*, etc. (BENJAMINI, 1988).

Os gráficos têm a função de informação, e proporcionam uma visualização rápida dos dados, descrevendo suas características e fornecendo elementos úteis para uma análise dos dados (PEARSON, 2001). O gráfico de análise utilizado nos capítulos a seguir foi *Box-plot*.

O gráfico *Box-plot* identifica os pontos discrepantes, fornecendo o resumo dos cinco números (mínimo, 1º quartil, mediana ou 2º quartil, 3º quartil e o valor máximo dos dados), permitindo comparar as variáveis da mesma natureza em uma mesma escala (BUSSAB e MORETTIN, 1987; TOLEDO e OVALLE, 1995), identificando características que possam evidenciar futuros problemas na validação dos pressupostos de uma metodologia de análise de dados (CHOONPRADUB e MCNEIL, 2005; PEARSON, 2001).

7 ESTATÍSTICA NÃO PARAMÉTRICA.

Estatística não-paramétrica pode ser definida como uma coleção de métodos estatísticos aplicada a conjunto de dados onde as suposições distribucionais necessárias para aplicação de uma técnica clássica (Intervalo de Confiança, Teste de Hipótese) não são satisfatoriamente atendidas, dispensando normalidade dos dados, o p-valor é exato, no caso paramétrico o cálculo do p-valor baseia na distribuição Normal (PONTES, 2005).

Os métodos não-paramétricos, em muitos casos não requerem cálculos complexos, são úteis quando é difícil estabelecer uma escala de valores quantitativa para os dados e mais eficientes que os paramétricos quando não existe normalidade. Entretanto, há desperdício de informações, já que em geral não consideram a magnitude dos dados. Quando as suposições

do modelo estatístico são atendidas, estes são menos eficientes que os paramétricos, porém, além do mais a utilização das tabelas dos métodos é mais complicada (SIEGEL, 1975).

8 TESTES NÃO PARAMÉTRICOS

Os testes paramétricos são aplicados conforme a característica e o número de amostras e o nível de associação entre duas ou mais amostras (GUIMARÃES, 2005)..

- **Teste para o caso de uma amostra.** Estes testes são úteis para verificar se determinada amostra pode provir de uma população especificada. São usualmente conhecidos como testes de aderência ou bondade do ajuste. Aonde se retira uma amostra aleatória e compara-se a distribuição amostral empírica com a distribuição teórica de interesse.
- **Teste para o caso de duas amostras relacionadas.** Este teste é aplicado quando se tem interesse na comparação de dois grupos tratamentos. Submetem-se dois tratamentos a um individuo em ocasiões diferentes. Outra maneira para pareamento é tentar selecionar indivíduos semelhantes como gêmeos, órgãos (ouvidos, braços, pés, etc...).
- **Teste para o caso de duas amostras independentes.** Testes que se aplicam aos planos amostrais onde se deseja comparar dois grupos independentes. Esses grupos podem ter sido formados de duas maneiras diferentes.
 - a) Extraiu-se uma amostra da população A e outra amostra da população B;
 - b) Indivíduos da mesma população foram alocados aleatoriamente a um dos dois tratamentos em estudo.
- **Teste para o caso de K Amostras relacionadas.** Técnica não-paramétrica de comparação de 3 ou mais grupos relacionados entre si (SIEGEL, 1975).

Diferente de dados pareados, não se exige que as amostra tenham o mesmo tamanho, ou seja, a idéia é comparação de médias dos grupos.

Algumas das vantagens na aplicação dos testes não-paramétricos é o fato de serem úteis em amostras pequenas e não haver necessidade dos dados corroborarem com distribuição probabilística normal e as variâncias poderem ser heterogenia. Portanto não há necessidade dos dados atenderem os pressupostos exigidos pelo modelo matemático, podendo os testes não paramétricos ser aplicados a dados que não sejam numéricos. Também é

possível ser utilizado quando é difícil estabelecer uma verdadeira escala quantitativa dos dados. Outro, porém, a quantidade de cálculos inferior aos dos testes paramétricos.

Existem testes não-paramétricos para lidar com amostras de populações com distribuição diferentes, nenhum teste paramétrico tem este poder (GOMES, 1990).

Entretanto as desvantagens do teste não-paramétrico é ter menor poder que os paramétricos, quando a população pode ser definida na sua forma de distribuição. Não são muito eficientes para rejeitar uma hipótese nula. Outro, porém, não leva em consideração a magnitude precisa dos dados.

9 TESTE DE HIPÓTESE

É uma regra de decisão que, com base na amostra, irá rejeitar ou não a hipótese nula (H_0). Hipótese nula é uma afirmação acerca do parâmetro (ou parâmetros) em questão que expressa sempre uma posição conservadora, em forma de igualdade, que se deseja sempre rejeitar, por exemplo, testar se os efeitos dos tratamentos aplicados são iguais a zero (GUIMARÃES, 2005). A afirmação oposta a H_0 chama-se hipótese alternativa (H_a), nela, por exemplo, deseja-se testar se pelo menos o efeito de um dos tratamentos aplicado difere de zero, no caso, deseja-se encontrar diferenças significativas.

Para verificar a significância do teste para rejeitar ou não H_0 , é necessário verificar se o p-valor está na região crítica para rejeitar H_0 . A região crítica é um subconjunto de valores tão extremos a curva da distribuição Normal, que nos leva a rejeição da H_0 quando o valor da estatística está contido nestes extremos (GUIMARÃES, 2005).

10 MEDIDAS DE CORRELAÇÃO.

Quando objetivo é verificar a existência de associação entre dois conjuntos de escores e também o grau de associação, utilizam-se medidas de correlação, conhecidas como Coeficiente de contingência C, Coeficiente de correlação de Spearman, Pearson, Kendall e a correlação canônica (ESTELBERGER e REIBNEGGER, 1995).

Existem várias formas de analisar a associação entre uma variável resposta com uma ou mais variáveis explicativas. O modelo mais comum de análise é a regressão múltipla. Na regressão múltipla, uma variável é explicada por uma combinação linear de outras variáveis (ABREU e VETTER, 1978).

A variável a ser explicada é chamada de dependente e as variáveis explicativas são chamadas de independentes.

COSTA (1977) sugerem que uma alternativa de analisar os dados é verificando a dependência entre as variáveis respostas e as variáveis explicativas, determinando o grau de associação entre duas variáveis por meio de coeficientes que apresentam propriedades adimensionais variando de -1 até 1. Os valores positivos destes coeficientes indicam uma correlação linear positiva, os valores negativos, uma correlação linear negativa e os valores nulos ou próximos de zero indicam ausência de correlação linear.

Para todos os coeficientes de correlação é necessário realizar a prova de significância para verificar o poder do teste (GIOLO, 2005). Entretanto, é freqüente a utilização dos testes de correlação sem a verificação do poder do teste.

Outro, porém, é verificar associação entre as variáveis por meio de gráficos, para identificar associação não linear entre as variáveis.

10.1 ANÁLISE DE CONTINGÊNCIA C

O Coeficiente de Contingência C tem objetivo de medir a associação entre dois conjuntos de atributos quando um ou ambos os conjuntos são medidos em escala nominal (GIOLO, 2005).

10.2 ANÁLISE DE CORRELAÇÃO DE SPEARMAN

O Coeficiente de Correlação de Spearman tem objetivo de medir a associação entre dois conjuntos de atributos quando ambos os conjuntos são medidos em escala de mensuração pelo menos ordinal, aonde as variáveis em estudo possam ser classificadas por postos em duas séries ordenadas. Têm-se n indivíduos “ordenados” em postos segundo duas variáveis X e Y (CAVADIAS, PILON e YUE, 2002). Para o cálculo da correlação de Spearman entre o conjunto de dados da variável X_i com a Y_i é necessário calcular a utilizando às seguintes formulas:

$$d_i = X_i - Y_i$$

$$r_s = 1 - \frac{6 \sum d_i^2}{(n^3 - n)}$$

Aonde d_i 's são diferenças entre os pares ordenados dos conjuntos e n é o tamanho da amostra.

Porém observado dois ou mais indivíduos com mesmo escore, é atribuído a cada um a média dos postos que lhe caberiam caso não tivesse havido empate. Em grandes quantidades de empates utiliza-se o seguinte fator de correção para r_s :

$$r_s = 1 - \frac{x^2 + y^2 + d}{2 * (x^2 + y^2)^{1/2}}$$

com

$$x^2 = \frac{n^3 - n}{12} - T_x \quad \text{em que } T_x = (t^3 - t)/12 \text{ e } t = \text{n}^\circ \text{ de observações empatadas em cada}$$

grupo de empates na variável X, e

$$y^2 = \frac{n^3 - n}{12} - T_y \quad \text{em que } T_y = (t^3 - t)/12 \text{ e } t = \text{n}^\circ \text{ de observações empatadas em cada}$$

grupo de empates na variável Y.

Entretanto, para fundamentar as conclusões baseadas nos resultados dos métodos de correlação é necessário verificar o nível de significância estatística do coeficiente (GIOLO, 2005).

Para realizar a prova de significância de r_s para comprovar a hipótese nula H_0 é necessário para amostra entre 4 a 30 observações utilizar a tabela P de Siegel, para obter a probabilidade associada à ocorrência de valores tão extremos quanto observado de r_s . Para amostra maiores que 10 observações utiliza-se a seguinte equação baseada na distribuição de t de Student com n-2 graus de liberdade:

$$t = r_s \sqrt{\frac{(n-2)}{1-r_s^2}}$$

10.3 ANÁLISE DE CORRELAÇÃO DE PEARSON

Equivalente ao coeficiente de correlação de Spearman tem-se o coeficiente de correlação de Pearson, podendo ser aplicado somente as variáveis contínuas ordenadas, desejando verificar a existência de correlação linear entre duas variáveis X e Y (BEST e RAYNER, 1996; GIOLO, 2005).

O coeficiente de correlação de Pearson (r) é dado pela seguinte equação:

$$r = \frac{1}{n-1} \frac{\sum_i (X_i - \bar{x})(Y_i - \bar{y})}{S_x S_y}$$

Em que S_x = desvio padrão amostral de X e S_y = desvio padrão amostral de Y.

Teste de significância do coeficiente é dado por:

$$t = r \sqrt{\frac{(n-2)}{1-r^2}} \sim t\text{- Studente com } n-2 \text{ g.l.}$$

A não rejeição da H_0 evidência que não há correlação linear entre as variáveis, podendo ocorrer uma associação não linear entre as variáveis, ou não haver associação.

10.4 ANÁLISE DE CORRELAÇÃO CANÔNICA

A técnica de correlação canônica pode ser muito útil em problemas que possuam mais de uma variável métrica dependente. O uso da correlação canônica pode simplificar o problema e determinar quais variáveis são mais importantes na análise. Desta forma, pode-se realizar a análise em duas etapas, primeiro determinando os fatores relevantes, e posteriormente realizando regressões simples entre os mesmos.

Pode haver uma significativa redução de complexidade e conseqüente facilidade na interpretação do problema. Segundo ABREU e VETTER (1978) a análise de correlação canônica é uma extensão da regressão múltipla, aonde o interesse é verificar a relação de duas ou mais variáveis dependentes, ou seja, desenvolver uma combinação linear em cada um dos conjuntos de variáveis, tal que a correlação entre os dois conjuntos seja maximizada. A equação é expressa por:

$$Y_1 + Y_2 + Y_3 + \dots + Y_n = X_1 + X_2 + X_3 + \dots + X_n$$

Uma forma de expressar uma correlação canônica pode ser:

$$V(X_1) = E\{(X_1 - \mu_1)(X_1 - \mu_1)'\} = \Sigma_1$$

$$V(X_2) = E\{(X_2 - \mu_2)(X_2 - \mu_2)'\} = \Sigma_2$$

$$COV(X_1, X_2) = E\{(X_1 - \mu_1)(X_2 - \mu_2)'\} = \Sigma_{12} = \Sigma_{21},$$

sendo assim uma matriz de covariância cruzada.

Supondo X_1 sendo uma matriz $n \times p$ e X_2 $n \times q$, tem-se a matriz $\text{COV}(X_1, X_2)$ em quatro partes:

$$\text{COV}(X_1, X_2) = \begin{matrix} & \begin{matrix} 11 & 12 \\ p \times p & p \times q \end{matrix} \\ \begin{matrix} 21 \\ q \times p \end{matrix} & \begin{matrix} 22 \\ q \times q \end{matrix} \end{matrix}$$

As covariâncias entre variáveis de diferentes conjuntos, uma variável de X_1 e outra de X_2 estarão contidas em (12) ou (21). Analisar as covariâncias em (12) ou (21) pode ser extremamente trabalhoso, ainda mais se p e q forem grandes. Porém, o principal objetivo da correlação canônica é resumir as associações entre X_1 e X_2 em função de algumas poucas correlações escolhidas, ao invés das pxq correlações.

Combinação linear é uma forma simples de resumir um conjunto de variáveis, então seja: $U = a'X_1$; $V = b'X_2$.

$$\begin{aligned} \text{Var}(U) &= a' \text{Cov}(X_1) a = a' \begin{matrix} 11 \\ p \times p \end{matrix} a \\ \text{Var}(V) &= b' \text{Cov}(X_2) b = b' \begin{matrix} 22 \\ q \times q \end{matrix} b \\ \text{Cov}(U, V) &= a' \text{Cov}(X_1, X_2) b = a' \begin{matrix} 12 \\ p \times q \end{matrix} b \end{aligned}$$

O que a correlação canônica procura é determinar os vetores a e b que maximizem a função.

$$\text{Corr}(U, V) = \frac{a' \begin{matrix} 12 \\ p \times q \end{matrix} b}{\sqrt{a' \begin{matrix} 11 \\ p \times p \end{matrix} a} \sqrt{b' \begin{matrix} 22 \\ q \times q \end{matrix} b}}$$

Existirão mínimo $(p, q) - 1$ pares de variáveis canônicas independentes do par de correlação máxima, que irão expressar a variância total dos dois grupos de variáveis.

Para calcular este máximo,

$$\max_{a, b} \text{Corr}(U, V) = \rho_1^*$$

restrita pela combinação linear (primeiro par de variáveis canônicas):

$$U_1 = \underbrace{e_1' \begin{matrix} -1/2 \\ 11 \end{matrix}}_{a_1'} X_1 \quad e \quad V_1 = \underbrace{f_1' \begin{matrix} -1/2 \\ 22 \end{matrix}}_{b_1'} X_2$$

$$\rho_1^{*2} \quad \begin{matrix} -1/2 & -1 \\ 11 & 12 & 22 & 21 & -1/2 \end{matrix} \quad e_1$$

Neste caso, é o autovalor de e é seu respectivo autovetor.

Supondo 'A' uma matriz quadrada kxk e 'v' um vetor kx1, pode-se mostrar que a equação $Av = \lambda v$ tem k soluções, sendo λ um escalar. Cada solução é dada por um par formado por um escalar λ_i e um vetor v_i . O escalar é chamado de autovalor de A e o vetor de autovetor de A.

A prova para o resultado acima pode ser encontrada em JOHNSON e WICHERN (1992).

O k-ésimo par de variáveis canônicas pode ser descrito como:

$$U_k = e'_k \begin{matrix} -1/2 \\ 11 \end{matrix} X_1 \quad e \quad V_k = f'_k \begin{matrix} -1/2 \\ 22 \end{matrix} X_2$$

maximizando,

$$\text{Corr}(U_k, V_k) \rho_k$$

Ambas as funções devem ser testadas, tanto isoladamente, como em conjunto. Os métodos a serem empregados são: lambda de Wilks, critério de Pillai, traço de Hotelling e a maior raiz de Roy. Todos estes métodos verificam a significância das funções do ponto de vista estatístico.

No mesmo sentido de realizar os testes de significância do coeficiente de Spearman e Pearson, pode ser utilizada a distribuição de qui-quadrado, conforme equação:

$$\chi^2 = -[N - 1 - 0,5(p + q + 1)] \times \ln \prod_{i=1}^q (1 - R_{ci}^2)$$

onde,

N = tamanho da amostra,

p = número de variáveis dependentes,

q = número de variáveis independentes,

R^2 = correlação canônica ao quadrado da equação a ser testada.

A distribuição de Qui-Quadrado em questão terá p x q graus de liberdade. A correlação canônica elevada ao quadrado é uma boa estimativa da variância compartilhada entre os dois grupos. O problema é que esta variação se refere somente a cada uma das

funções separadamente, fazendo com que uma parte da variância não seja levada em conta. Um novo índice pode ser calculado para evitar este viés, o índice de redundância.

Este índice pode ser calculado como a média dos quadrados dos coeficientes de correlação entre o total das variáveis independentes e cada variável dependente. Este índice corresponde a um resumo da habilidade de um conjunto de variáveis independentes (tomadas em conjunto) explicarem uma variação nas variáveis dependentes (tomadas uma a uma). Este índice pode ser comparado ao R^2 de uma regressão múltipla (NETO, 2005).

11 ANÁLISE DE VARIÂNCIA

Para análise de dados com medidas repetidas, existe uma grande variedade de técnicas quando a variável resposta segue uma distribuição normal. Entre elas, podemos destacar a análise de variância com medidas repetidas (NETER et. al.,1996), análises uni/multivariada de perfis, análises de curvas de crescimento e modelos normais com efeitos aleatórios.

A análise de Variância requer a verificação da validade de várias suposições fundamentais. Para cada nível verificar se os valores seguem uma distribuição Normal e as variâncias são as mesmas, bem como aditividade e independência dos resíduos do modelo (ROHATGI, 1976; AGUIRRE, 1997). É comum analisar graficamente os pressupostos (RYAN e JOINER, 1976), porém usuários sem experiências têm dificuldades na interpretação dos gráficos. Portanto, métodos estatísticos como Bartlett e Shapiro Wilk, podem ser obtidos facilmente por *softwares* específicos em análises estatísticas.

Os delineamentos experimentais têm objetivo de reduzir o erro, designando os tratamentos às unidades experimentais ou parcelas (COCHRAN e COX, 1978). Há diversos tipos de delineamentos tais como fatorial, hierárquico, hierárquico cruzado, em faixas, fracionais, quadrado latino, parcelas sub-dividida e etc.,(ANJOS, 2004).

11.1 PARCELAS SUB-SUB-DIVIDIDAS

Em experimentos fatoriais com dois ou mais fatores, quando há alguma limitação para instalar o experimento, ou mesmo para facilitar a instalação, utiliza-se o experimento em parcelas subdivididas. Em alguns casos, é a única forma de aplicação dos tratamentos às unidades experimentais (ANJOS, 2004).

Para analisar os dados por meio do delineamento de parcela sub-subdividida, além de verificar os pressupostos e características da variável, se faz necessário definir os fatores, blocos (repetição (A)), tratamentos da parcela principal (B) , subparcela (C) e sub-subparcela

(D). É uma extensão do delineamento de parcela subdividida. Os níveis de um fator são alocados nas unidades principais dentro de cada bloco.

Os níveis do segundo fator são alocados nas subparcelas dentro de cada parcela e as doses de um terceiro fator são alocados nas sub-subparcelas dentro de cada subparcela. Em geral, o fator de maior interesse é colocado nas subparcelas, quando possível. Caso contrário, a aplicação dos tratamentos às parcelas principais ou subparcelas, dependerá da facilidade de instalação do experimento.

Se existirem apenas dois fatores, o experimento é chamado de parcelas subdivididas.

11.2 MODELO MATEMÁTICO

Para um experimento em Parcelas Sub-subdivididas, o modelo estatístico (com blocos) e dado da seguinte forma:

$$y_{ijkl} = \mu + \rho_i + \alpha_j + \varepsilon_{ij} + \beta_k + \tau_l + (\alpha\beta)_{jk} + \varepsilon_{ijk} + \tau_l + (\alpha\tau)_{jl} + (\beta\tau)_{kl} + (\alpha\beta\tau)_{jkl} + \varepsilon_{ijkl}$$

μ = constante;

ρ_i = efeito do i-ésimo bloco $i = 1, \dots, r$;

α_j = efeito do j-ésimo nível do fator A (parcela principal) $j = 1, \dots, a$;

ε_{ij} = Resíduo a (interação entre blocos ou repetições e o fator A ou Parcela Principal)

$$\varepsilon_{ij} \stackrel{iid}{\sim} N(0, \sigma_\gamma^2);$$

β_k = efeito do k-ésimo nível do fator B (Subparcela) $k = 1, \dots, b$;

$(\alpha\beta)_{jk}$ = efeito da interação entre j-ésimo nível do fator A e k-ésimo nível do fator B;

ε_{jk} = Resíduo b (interação entre A Parcela Principal e o fator B ou Parcela Subparcela)

$$\varepsilon_{jk} \stackrel{iid}{\sim} N(0, \sigma_\gamma^2)$$

τ_l = efeito do l-ésimo nível do fator C (Sub-subparcela) $k = 1, \dots, c$;

$(\alpha\tau)_{jl}$ = efeito da interação entre j-ésimo nível do fator A e l-ésimo nível do fator C;

$(\beta\tau)_{kl}$ = efeito da interação entre k-ésimo nível do fator B e l-ésimo nível do fator C;

$(\alpha\beta\tau)_{jkl}$ = efeito da interação entre o j-ésimo nível do fator A, K-ésimo nível do fator B e l-ésimo nível do fator C;

\mathcal{E}_{ijkl} = resíduo c, associado interação do fator B (sub-parcela) e C (sub-subparcela);

$$\mathcal{E}_{ijkl} \stackrel{iid}{\sim} N(0, \sigma_{\gamma}^2)$$

11.3 DECOMPOSIÇÃO DAS CAUSAS DE VARIAÇÃO

No Quadro 1 é possível verificar a decomposição da variação total das observações em partes que podem ser atribuídas a fatores conhecidos e controláveis e em fatores desconhecidos e não controláveis, ou seja o erro ou resíduo. Os erros podem ser atribuídos a forma de condução do experimento, o ambiente em que se está conduzindo ou material utilizado no experimento.

QUADRO 1 – DESCRIÇÃO DA ANÁLISE DE VARIÂNCIA PARA O DELINEAMENTO DE PARCELA SUBSUBDIVIDIDA.

CAUSA DE VARIAÇÃO	GL	QM	F
A (Blocos)	(a-1)	QMA	
B	(b-1)	QMB	QMB/QME(a)
Erro (a)	(a-1)(b-1)	QME(a)	
PARCELA			
C	(c-1)	QMC	QMC/QME(b)
BxC	(b-1)(c-1)	QM(BxC)	QM(BxC)/QME(b)
Erro (b)	b(c-1)(a-1)	QME(b)	
SUB-PARCELAS			
D	(d-1)	QMD	QMD/QME(c)
BD	(b-1)(d-1)	QM(BxD)	QM(BxD)/QME(c)
CD	(c-1)(d-1)	QM(CxD)	QM(CxD)/QME(c)
BCD	(b-1)(c-1)(d-1)	QM(BxCxD)	QM(BxCxD)/QME(c)
Erro (c)	BC(d-1)(a-1)	QME(c)	
TOTAL	abcd-1		

A forma analítica de encontrar as esperanças dos quadrados médios $E(QM)$ requer cálculos extensos e muitas vezes complexos. HICKS (1973) citado por ANJOS (2004), desenvolveu um algoritmo para se encontrar $E(QM)$ que torna mais simples e menos dispendioso esta tarefa. Alguns *softwares* possuem o recurso de definição do tipo de efeito do fator e automaticamente realizam o teste F de maneira correta. No caso específico do *software* R (R *DEVELOPMENT CORE TEAM.*), esta definição deve ser feita para todos os casos em que o denominador (ou termo de erro) não seja o Quadrado Médio do Resíduo. Deve-se tomar

cuidado neste ponto, pois em geral, tanto o R como outros *softwares* não mencionam qualquer fato a respeito de problemas desta natureza. Em resumo, o resultado da análise é fornecido mas, pode estar parcialmente ou totalmente incorreto (ANJOS, 2004; MONTGOMERY, 2001).

11.4 SOMA DE QUADRADOS

Os cálculos para se obter a Soma de Quadrados (SQ) do experimento delineado com parcelas sub-subdividida pode ser obtida da seguinte forma:

$$SQ_{Subparcela} = \frac{\sum_{i,j,k} y_{ijk}^2}{d} - FC$$

$$SQD = \frac{\sum_{...l} y_{...l}^2}{abc} - FC$$

$$SQ_{CD} = \frac{\sum_{j,l} y_{.jl}^2}{ab} - FC - SQ_C - SQD$$

$$SQ_{BD} = \frac{\sum_{j,l} y_{.jl}^2}{ac} - FC - SQ_B - SQD$$

$$SQ_{BCD} = \frac{\sum_{j,k,l} y_{j.k.l}^2}{a} - FC - SQ_B - SQD - SQ_{BC} - SQ_{BD} - SQ_{CD}$$

$$SQE(c) = SQ_{Total} - SQ_{Subparcela} - SQ_{PP} - SQD - SQ_{BD} - SQ_{CD} - SQ_{BCD}$$

$$SQ_{Total} = \sum_{i,j,k,l} y_{ijkl}^2 - FC$$

No entanto, atualmente não é necessário aplicar estas fórmulas para obter o quadro da ANOVA, os cálculos já estão implementados no *software* R.

É comum no setor agrário que os estudos fundamentem as análises utilizando delineamentos em parcelas sub-subdivididas. No entanto, algumas análises de variância foram realizadas para variáveis mensuradas por meio de contagem, portanto, observou-se perda de poder do Teste F. Outro, porém a ser considerando é a dependência temporal, como é

identificada no estudo realizado por MURARO (2004), com épocas de deferimento, adubação nitrogenada e pastejo, afetando parâmetros radiculares de mistura de aveia preta e azevém, utilizou ANOVA somente para variáveis contínuas, no entanto em dado momento não foi verificado a dependência temporal nas épocas de deferimento. Portanto é comum utilizar ANOVA em pesquisas agrárias, porém a verificação dos pressupostos para utilização da metodologia de análise por vezes não é checada (SCHWERTMAN e SILVA, 2007).

No capítulo 2, realizou-se análises ANOVA, partindo do pressuposto que há independência entre valores mensurados por um determinado tempo e a análise de modelos lineares generalizados para dados longitudinais e por fim, uma análise não paramétrica por meio de árvore de classificação.

11.5 TESTE DE COMPARAÇÃO DE MÉDIAS

Ao concluir que existe diferença significativa entre tratamentos através do teste F (ANOVA), avaliar a magnitude destas diferenças por meio de um teste de comparações múltiplas. Dentre os diversos testes, optou-se pelo Teste de Tukey freqüentemente em análises de experimentos agrícolas, este permite testar qualquer contraste, sempre, entre duas médias de tratamentos, ou seja, não permite comparar grupos entre si. O teste baseia-se na Diferença Mínima Significativa (DMS) Δ . A estatística do teste é dada da seguinte forma:

$$\Delta = q \sqrt{\frac{QMRES}{r}}$$

onde q é a amplitude total studentizada, tabelada, QMRES é o quadrado médio do resíduo e r o número de repetições. O valor de q depende do número de tratamentos e do número de graus de liberdade do resíduo.

Em geral, é comum em experimentos agrícolas o nível de confiança adotado ser de 95%.

Como o teste de Tukey é independente do teste F, é possível que, mesmo sendo significativo o valor de F calculado, não se encontre diferenças significativas entre contrastes de médias.

Para concluir que o contraste é significativo, a diferença entre médias deve ser maior do que o valor de Δ . É comum para apresentar o resultados do teste de Tukey, utilizar o método de letras na tabela.

11.6 TRANSFORMAÇÃO BOX-COX

Há diversas formas de realizar transformação nos dados para atender pressuposto da ANOVA, a que tem apresentado melhores resultados é a transformação Box-Cox, que consiste em transformar os dados de acordo com a expressão

$$y^* = y^\lambda$$

onde λ é um parâmetro a ser estimado dos dados. Se $\lambda = 0$ a equação acima se reduz a

$$y^* = \log(y)$$

onde Log é o logaritmo neperiano. Uma vez obtido o valor de λ encontra-se os valores dos dados transformados conforme a equação acima e utilizam-se estes dados transformados para efetuar as análises.

No *software* R a função Box-cox do pacote MASS calcula a verossimilhança perfilhada do parâmetro λ . Escolhe-se o valor que maximiza esta função. É possível visualizar o valor no gráfico da verossimilhança perfilhada.

12 MODELOS LINEARES GENERALIZADOS

Outra opção de análise dos dados, quando não é possível utilizar ANOVA em função dos pressupostos não estarem sendo atendidos, é Modelos Lineares Generalizados. Este que possibilita determinar a melhor distribuição teórica que represente os dados, ao contrário da ANOVA que somente é possível aplicação a dados que seguem uma distribuição Normal.

O estudo realizado por PÉREZ e FREIRE (2003), trata da evolução do número de empregados em empresas, ao longo de um certo período de tempo, segundo alguma determinada classificação, aonde, o número de empregados é a variável resposta, o qual conclui-se que a Análise de Variância não foi apropriada para subsidiar tomadas de decisões, em função da característica da variável reposta.

Nas Ciências Agrárias utiliza-se com freqüência a ANOVA, fundamentada na distribuição normal (NETER et al.,1996) para dados de proporção e contagem. No entanto, segundo a distribuição probabilística geralmente utilizada para análise de dados de proporções é a Binomial e para dados de contagem à distribuição de Poisson (PÉREZ e FREIRE; 2003), sendo que ambas são casos particulares do Modelos Lineares Generalizados (NELDER e WEDDERBURN, 1972).

Segundo NELDER e WEDDERBURN, (1972) citado por COSTA (2003), a utilização de modelos lineares clássicos não é, em geral, apropriada para analisar dados de proporções,

por consequência dos pressupostos do modelo não serem atendidos. Portanto a teoria de Modelos Lineares Generalizados (GLM) é o método mais indicado para analisar dados que não atendem os pressupostos de normalidade e homocedasticidade dos resíduos.

No capítulo 2, estudou-se em Campo Largo e Lapa, durante 2 e 3 safras respectivamente, o nível de ácaros rajados, em pomares de pessegueiro, sobre diferentes níveis de adubação nitrogenada e potássica. Neste caso, segundo LIANG e ZEGER (1986), estudos longitudinais são de particular interesse quando objetivo é avaliar variações globais ou individuais ao longo do tempo. Sendo que para variável resposta sendo número de ácaros na planta, a distribuição Probabilística de Poisson é a mais indicada para realizar as análises de comparação entre os tratamentos realizados no solo (PÉREZ e FREIRE, 2003). Já no capítulo 3, devido as características das variáveis respostas serem contínuas, dependendo dos pressupostos, é possível realizar as análises tanto pela ANOVA como por GLM, ambas baseadas na distribuição Normal (MCCULLAGH e NELDER, 1989; DEMETRIO, 2001).

12.1 DEFINIÇÃO

Segundo MCCULLAGH e NELDER (1989) o modelo linear generalizado se caracteriza da seguinte forma: suponha que Y_1, \dots, Y_n são variáveis aleatórias independentes, de média μ_i , cada uma com função de probabilidade em densidade ou função densidade de probabilidade pertencente à família exponencial apresentada na seguinte forma:

$$f(y_i | \theta_i, \phi) = \exp\{[y_i \theta_i - b(\theta_i) + c(y_i) + a(y_i, \phi)]\},$$

onde $E\{Y_i\} = db(\theta_i)/d\theta_i$, que é denotado por μ_i , $\text{var}\{Y_i\} = \phi^{-1} V_i V_i = d\mu_i/d\theta_i$ é a função de variância, $\theta = \theta(\beta)$ é o parâmetro canônico e ϕ é o parâmetro de dispersão ($\phi > 0$), geralmente desconhecido ($i = 1, 2, \dots, n$).

Caso ϕ seja conhecido tem-se uma distribuição da família exponencial, caso contrário poderá pertencer a outra família.

A função $a_i(\phi)$ geralmente assume a seguinte forma:

$$a_i(\phi) = \frac{\phi}{\omega_i}$$

onde se assume que ω_i é conhecido, enquanto que ϕ_i , podendo ser denominado parâmetros de dispersão ou escalar, portanto contante, ou seja, não varia com o índice i das variáveis aleatórias y_i .

Os modelos lineares generalizados são definidos então pelo componente sistemático

$$g(\mu_i) = g(E[Y_i]) = x_i^T \beta$$

onde:

x_i = vetor de constantes conhecidas para o i -ésimo sujeito.

β = vetor de parâmetros de regressão.

$g(\cdot)$ = função monótona e diferenciável, denominada função de ligação.

De acordo com DRAPER e SMITH (1998) a variância em modelos lineares generalizados pode ser expressa da seguinte forma:

$$\sigma_i^2 = \phi_i V(\mu_i),$$

e o parâmetro de dispersão pode ser modelado por:

$$h(\phi_i) = z_i^T \gamma,$$

onde:

$h(\cdot)$ = outra função de ligação

z_i = outro vetor de constantes conhecidas para o i -ésimo sujeito.

γ = outro vetor de parâmetros de regressão (explicando o comportamento da variância).

Em um modelo marginal o interesse está em modelar a esperança marginal ou a resposta média marginal em função de determinadas variáveis explanatórias. Considerando as definições anteriores os modelos marginais de regressão serão:

Modelo para a média: $g(\mu_{ij}) = g(E[Y_{ij}]) = x_{ij}^T \beta$

Modelo para a dispersão: $h(\phi_{ij}) = z_{ij}^T \gamma$

Portanto, de acordo com a natureza das variáveis resposta, pode-se utilizar diferentes modelos marginais, como por exemplo, a Normal, Poisson e Binomial tem-se respectivamente os seguintes modelos:

$$\mu = \eta ; \quad \log \mu = \eta \quad \text{e} \quad \log \frac{\mu}{1 - \mu} = \beta_0 + \sum_{j=1}^p \beta_j x_j$$

Apresenta-se a seguir algumas distribuições particulares da família exponencial.

Em casos particulares a binomial em que Y^* é a proporção de sucessos em n ensaios independentes, cada um com probabilidade de ocorrência μ . Assumindo $nY^* \sim B(n, \mu)$. A densidade de Y^* é expressada na forma:

$$\frac{n}{ny^*} \mu^{ny^*} (1 - \mu)^{n - ny^*} = \exp \left[\log \frac{n}{ny^*} + ny^* \log \frac{\mu}{1 - \mu} + n \log (1 - \mu) \right],$$

onde, $0 < \mu, y^* < 1$.

Obtêm-se a forma da função densidade da família exponencial fazendo:

$$\phi = 0, \theta = \log\{\mu/(1 - \mu)\}, b(\theta) = \log(1 + e^\theta), c(y^*, \phi) = \log \frac{\phi}{\phi y^*};$$

a função da variância fica: $V(\mu) = \mu(1 - \mu)$

12.2 LIGAÇÕES CANÔNICAS

As ligações canônicas são obtidas pelo logaritmo da função de verossimilhança de um MLG com respostas independentes é expresso da seguinte forma:

$$L(\beta; y) = \sum_{i=1}^n \phi \{y_i \theta_i - b(\theta_i)\} + \sum_{i=1}^n c(y_i, \phi)$$

Um caso particular importante ocorre quando o parâmetro canônico (θ) coincide com o preditor linear, ou seja, quando $\theta_i = \eta_i = \sum_{j=1}^p x_{ij} \beta_j$. Nesse sentido, $L(\beta; y)$ fica dado por:

$$L(\beta; y) = \sum_{i=1}^n \phi \sum_{j=1}^p x_{ij} \beta_j - b \sum_{j=1}^p x_{ij} \beta_j + \sum_{i=1}^n c(y_i, \phi)$$

Definindo a estatística $S_j = \phi \sum_{i=1}^n Y_i x_{ij}$, $L(\beta; y)$ fica então expresso na forma:

$$L(\beta; y) = \sum_{j=1}^p s_j \beta_j - \phi \sum_{i=1}^n b \sum_{j=1}^p x_{ij} \beta_j + \sum_{i=1}^n c(y_i, \phi)$$

Pelo teorema de fatorização a estatística $S = (S_1, \dots, S_p)^T$ é suficiente minimal para o vetor $\beta = (\beta_1, \dots, \beta_p)^T$. As ligações que correspondem a tais estatísticas são chamadas de ligações canônicas e desempenham um papel importante na teoria dos MLGs.

Uma das vantagens da utilização de ligações canônicas é que elas garantem a concavidade de $L(\beta; y)$ e conseqüentemente muitos resultados assintóticos são obtidos mais facilmente.

12.3 ESTRUTURA DE CORRELAÇÃO

Quando se trata de informações que foram coletadas ao longo do tempo torna-se necessário modelar uma estrutura de covariância entre as observações em um mesmo indivíduo, para tanto assume-se que:

$$\Sigma_i = \Sigma_i^{1/2} \Sigma_i \Sigma_i^{1/2}$$

12.4 TESTES DE HIPÓTESES

Os métodos de inferência nos modelos lineares generalizados são baseados na teoria de máxima verossimilhança. De acordo com esta teoria, existem três estatísticas para testar hipóteses relativas aos parâmetros β 's que são: (RUSSO, 2006).

- Razão de verossimilhanças
- Wald
- Escore

Essas estatísticas são assintoticamente equivalentes e, sob H_0 e para ϕ conhecido, convergem para uma variável com distribuição χ_p^2 .

Dentre os métodos citados, o teste de Wald, que é baseado na distribuição normal assintótica de $\hat{\beta}$ e é uma generalização da estatística t de Student (RUSSO, 2006). É geralmente o mais usado no caso de hipóteses relativas a um único coeficiente β_j . A estatística para este teste é:

$$W = (\hat{\beta}_1 - \beta_{1,0})^T \hat{Var}(\hat{\beta}_1)^{-1} (\hat{\beta}_1 - \beta_{1,0}),$$

sendo $\hat{Var}(\hat{\beta}_1)$ a $Var(\hat{\beta}_1)$ avaliada em $\hat{\beta} = \begin{pmatrix} \hat{\beta}_1^T \\ \hat{\beta}_2^T \end{pmatrix}$. Para amostras grandes, rejeita-se

H_0 , a um nível de $100\alpha\%$ de probabilidade, se $W > \chi_{q,1-\alpha}$

13 ANÁLISE DE CLASSIFICAÇÃO

Quando se trabalha com grandes quantidades de variáveis, a análise multivariada é a mais indicada para extrair informações da base de dados. As técnicas para classificação, estimação, predição, análise de agrupamentos, análise de associação são algumas análises multivariadas (NETO, 2005).

As ferramentas estatísticas ajudam a encontrar modelos matemáticos a partir de estruturas que representam o conjunto de dados. Porém, não há técnica que resolva todos os problemas, os métodos se diferenciam pelos seus objetivos, estes tendo vantagens e desvantagens (MASSAD e SILVEIRA, 2003).

13.1 REDES NEURAIS.

É uma classe de modelagem de prognóstico que trabalha com ajustes repetido de parâmetros, inspirados nos neurônios biológicos e na estrutura massivamente paralela do cérebro, exibindo características similares ao do comportamento humano. Estruturalmente, uma rede neural consiste em um número de elementos interconectados e organizados em camadas que aprendem pela modificação da conexão representada por equações complexas, sendo assim capazes de generalizar seu conhecimento a partir de exemplos anteriores (FAYYAD, 1996). Os elementos interconectados, chamados de neurônios, avaliam os valores de entrada, calculam o total para valores de entrada combinados, comparam o total com um valor limiar, e determinam o que sairá. A operação dos neurônios é simples, o que torna complexo o processo é o conjunto de operações realizadas pelos neurônios (FERREIRA e ALVES, 2005).

13.2 ANÁLISE DISCRIMINANTE

Análise discriminante tem como objetivo determinar a qual grupo entre dois ou mais definido a priori, pertence a um determinado elemento. A determinação é realizada considerando as características das variáveis aleatórias que contribui com informação para referida classificação (FAYYAD, 1996). A eficiência de uma técnica é proporcional à qualidade das informações disponíveis. Isto dá à fase de coleta de dados uma importância fundamental. Independente do método definido, se as variáveis forem selecionadas de forma inadequada acaba por comprometer a eficiência almejada. A Análise Discriminante combina variável em uma ou mais funções determinando os valores para a classificação. Estas funções são construídas de modo que os escores dos elementos de cada grupo se concentrem em torno do valor médio do grupo, fazendo com que a superposição de escores de elementos de diferentes grupos seja minimizada (HAIR, et. al., 1995).

13.3 ANÁLISE DE REGRESSÃO.

A regressão é utilizada para testar dependências cumulativas de uma variável, chamada de dependente, em relação a uma ou diversas variáveis independentes (GIOLO, 2005). A ferramenta de regressão pode ser classificada de três formas:

1. Regressão Linear Simples, que consiste em testar a relação de uma variável dependente uma variável independente;

2. Análise de Regressão Múltipla, que consiste em testar a relação de uma variável dependente com várias variáveis independentes;
3. Análise de Correlação Canônica, que consiste em testar a relação entre um conjunto de variáveis dependentes com outro conjunto de variáveis independentes.

As relações entre as variáveis dependentes e independentes podem ser representadas por diagramas de dispersão. Uma das principais aplicações da Regressão é identificar qual variável que mais contribui para formação da variável dependente (GIOLO, 2005).

13.4 REGRESSÃO LOGÍSTICA.

A regressão logística consiste na busca de um modelo que permita relacionar uma variável Y , chamada “variável resposta”, aos “fatores” X_1, \dots, X_{p-1} , que, supõe-se, influenciam as ocorrências de um evento (GIOLO, 2005).

A variável resposta deve ser do tipo dicotômica, assumindo apenas os valores 0 ou 1. Neste caso existe interesse apenas na ocorrência, ou não, do evento em questão. Quando se tem variável resposta classificada como dicotômica, não recomendam a aplicação de modelos lineares generalizados, pelos seguintes motivos: o Modelo Linear Geral pode gerar estimativas que não pertence ao intervalo $[0, 1]$, outro motivo é que a variância dos resíduos não é constante (GIOLO, 2005).

13.5 ANÁLISE DE CLUSTER

Análise de agrupamento, apesar de não ser um método de classificação é distinta dos demais métodos apresentados neste trabalho. Classificar é concernente com um número de grupos conhecidos e o objetivo operacional é fixar uma nova observação em um destes grupos. Agrupar é uma técnica mais primitiva no sentido de nenhuma suposição é feita quanto ao número de grupos ou estruturas de agrupamento. O agrupamento é realizado a com base na similaridade ou distâncias (NETO, 2005).

13.6

número relativamente pequeno destas variáveis, associando aos coeficientes que indica o quanto esta sendo explicado pela variável.

Outro ponto importante é o fato de que as funções empíricas derivadas desta técnica são favoráveis para interpretações físicas. Finalmente, a ACP é adequada para campos espaciais em grade. Portanto, os componentes principais são combinações lineares não correlacionadas, cujas variâncias são tão grandes quanto possíveis (CARVALHO et. al. 2003)

13.7 ÁRVORES DE CLASSIFICAÇÃO

Existem dois tipos de Árvores de Decisão: Árvore de Regressão, aplicadas a variáveis contínuas e Árvore de Classificação aplicada a variáveis discretas. Ambas contém nós com decisões preditivas. As árvores de Regressão são semelhantes às Árvores de Decisão. A diferença principal está nas folhas conterem previsões numéricas para Árvore de Regressão e não decisões como é para Árvore de Classificação.

13.7.1 INTERPRETAÇÃO

É uma técnica simples de expressar a lógica condicional. A Árvore de Decisão seleciona o atributo mais significativo aos dados, sendo um modelo preditivo representado por um gráfico em forma de árvore contendo as decisões a serem inferidas indutivamente (COELHO, 2005). Esta forma de árvore tem um conjunto de nós de decisão que se associam à covariável, a qual podem ser tomadas decisões depois de treinadas de acordo com um conjunto de treinos realizados na variável de interesse (CARVALHO, 2005). A decisão tomada é representada pela folha da árvore, representando o valor previsto da variável de interesse. A árvore pode ser analisada a partir do topo do gráfico também chamado nó raiz da árvore, aonde se localiza o fator mais expressivo (COELHO, 2005).

13.7.2 VANTAGENS

As vantagens ao analisar conjunto de dados por meio de árvores de classificação estão em: não requerer cálculos complexos; dados discrepantes não influenciam nas decisões, a falta de valores para alguns atributos não influencia nos valores preditivos; capacidade de aprender por meio de exemplos; facilidade na preparação dos dados para análises (CARVALHO, 2005).

13.7.3 DESVANTAGENS

As desvantagens são: pouca robustez a dados de grande dimensão, a precisão afetada é por atributos pouco relevantes e há dificuldade em lidar com dados contínuos (COELHO, 2005).

Ao analisar conjunto de dados utilizando árvores de regressão ou árvore de classificação se têm as mesmas vantagens e desvantagens, ficando a critério do pesquisador decidir sobre qual técnica é mais adequada aos dados (CARVALHO, 2005).

13.7.4 DEFINIÇÃO

Para desenvolver Árvore de Classificação, considera-se conjunto de dados Z , em que $Z = \{x_1, x_2, \dots, x_n\}$ e classes definidas como C , em que $C = \{c_1, c_2, \dots, c_n\}$ com $m < n$.

Tendo como objetivo estabelecer relação segundo uma função f definida como $f: Z \rightarrow C$, em que, a cada vetor x de Z (o espaço de medida), cujas coordenadas são os valores assumidos pelas variáveis explicativas que descrevem a amostra para o caso x , a uma classe de C .

Desta forma defini-se a regra de classificação como uma função de $f(x)$ definida em Z , que para todo o x pertencente a Z , $f(x)$ pertence a uma das classes c_1, c_2, \dots, c_n , ou equivalente, uma partição de Z em n subconjuntos disjuntos A_1, \dots, A_n com:

$$Z = \bigcup_{j \in c} A_j$$

tal que, para todo o x pertencente a A , a classe prevista é a j .

Porém, em geral, não estão disponível os dados da população, tamanho da amostra suficiente, desta forma dados de Z têm que ser usados para construir $f(x)$ e para estimar o valor do risco de classificação incorreta, $R^*(f)$, associado a f . Esta rotina é conhecida como validação interna. Considerando o método não-paramétrico de validação JACKKNIFE, destinara-se a estimar o enviesamento e a variância de estimadores em condições teoricamente complexas ou em que não se tem confiança no modelo especificado. É um método de reamostragem fundamentada em sub-amostragem com características semelhantes tanto em tamanho como na associação a uma distribuição da variável dependente. Na primeira etapa, uma das amostra é reservada para cálculo do erro da classificação, enquanto as demais serão utilizadas para criar as regras de classificação. Este processo repete-se até que cada uma das n_i amostras tenha sido utilizada para o cálculo do erro de classificação, o valor final deste

é obtido por meio da média associada as n_i amostras de precisão disponíveis (SCHWERTMAN e SILVA 2007).

14 BOOTSTRAPING

No estudo, capítulo 2, da classificação do número médio de ácaros na folha de pessegueiro em 3 níveis (baixo, médio e alto), foi possível após aplicação da metodologia *Bootstrapping*. Para aumentar o número de observações a partir da mesma amostra, em geral, a técnica *bootstrapping* é eficiente.

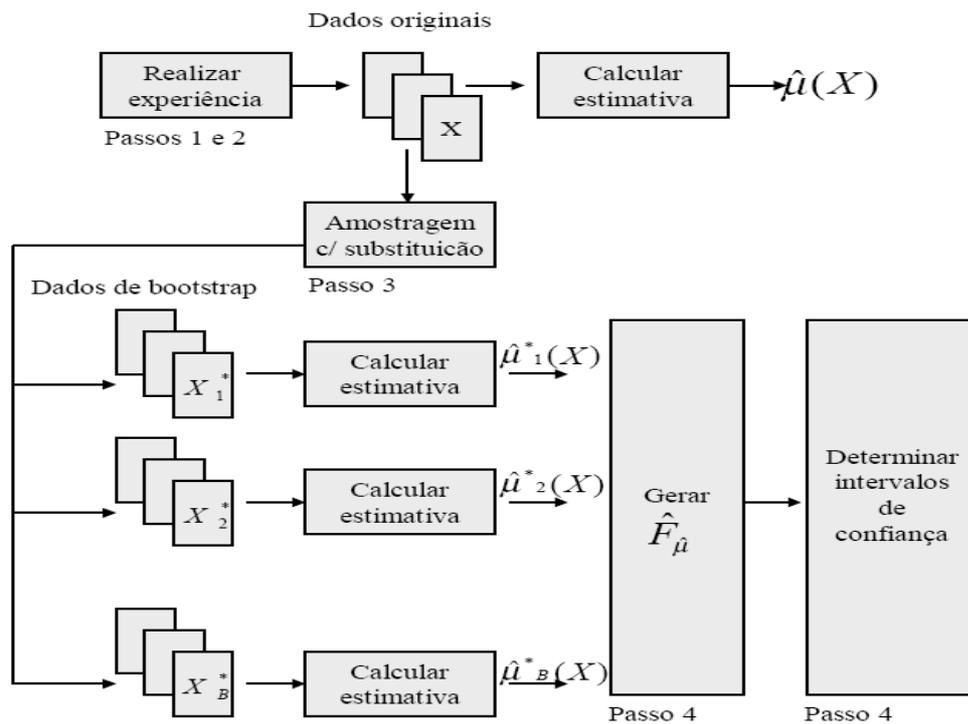


FIGURA 1 - PASSOS PARA APLICAÇÃO DA TÉCNICA *BOOTSTRAP*

Exemplificando: Seja a estimativa da média de uma variável aleatória X , com o seguinte conjunto de amostra:

$$X = (-2,41; 4,86; 6,06; 9,11; 10,2; 12,81; 13,17; 14,1; 15,77; 15,79)$$

Utilizando a formula abaixo para $n=10$ obtem-se

$$\hat{\mu} = \frac{X_1 + \dots + X_n}{N}$$

$$\hat{\mu} = 9,94$$

Selecionam-se aleatoriamente n amostras aleatoriamente com 10 observações e calcula-se um número elevado de vezes para obter as estimativas

$$\hat{\mu}_a, \dots, \hat{\mu}_b$$

$$\hat{\mu}_{(1)}, \hat{\mu}_{(2)}, \dots, \hat{\mu}_{(b)}$$

$\hat{\mu}_{(k)}$, é o k -ésimo valor mais pequeno de $\hat{\mu}_{(1)}, \dots, \hat{\mu}_{(b)}$

Para o intervalo de confiança a $(1-\alpha)$ 100% e dado por

$$\hat{\mu}_{(q_1)}, \dots, \hat{\mu}_{(q_2)}$$

em que

$q_1 = \text{parte inteira}(B\alpha/2)$

q_2

15.1 MÉTODOS DE CLASSIFICAÇÃO

No sistema WEKA pode ser encontrado os seguintes métodos de classificação:

- a) Árvore de Decisão Induzida;
- b) Regras de aprendizagem;
- c) *Naive Bayes*;
- d) Tabelas de decisão;
- e) Regressão local de pesos;
- f) Aprendizado baseado em instância;
- g) Regressão lógica;
- h) *Perceptron*;
- i) *Perceptron* multicamada;
- j) Comitê de *perceptrons*;
- k) *SVM*.

15.2 MÉTODOS DE PREDIÇÃO

Também são encontrados no WEKA os seguintes métodos para predição numérica:

- a) Regressão linear;
- b) Geradores de árvores modelo;
- c) Regressão local de pesos;
- d) Aprendizado baseado em instância;
- e) Tabelas de decisão;
- f) *Perceptron* multicamadas.

15.3 INICIALIZAÇÃO DAS ANÁLISES

Para inicialização da análise dos dados por meio do *software* WEKA, os dados podem estar armazenados em uma planilha do *Microsoft Excel* ou *Word*. Também é possível utilizar outros programas para converter os dados para o formato ARFF.

Para gerar um arquivo onde os dados serão separados por vírgula, deve clicar menu Arquivo seleciona a opção Salvar como e na caixa de diálogo Salvar como tipo a opção CSV.

O sistema Weka possui um formato próprio o ARFF, o qual é necessário descrever o domínio do atributo, pois o mesmo não pode ser obtido automaticamente pelo seu valor.

Antes da aplicação dos algoritmos do sistema WEKA, é necessário os dados estarem no formato ARFF que consiste em duas partes:

- Conter uma lista de todos os atributos, onde devemos definir o tipo do atributo ou os valores que ele pode representar, quando utilizamos os valores estes devem estar entre “{}” separados por vírgulas.
- Os registros (instâncias) a serem minerados com o valor dos atributos para cada instância separado por vírgula, a ausência de um item em um registro deve ser atribuída pelo símbolo “?”.

REFERÊNCIAS

- AGUIRRE, A. **Uma nota sobre a transformação BOX-COX**. Disponível em: <<http://www.cedeplar.ufmg.br/pesquisas/td/TD%20116.pdf>>. Acesso em Nov. 2007
- ABREU, M. A.; VETTER, D. A análise de relações entre conjuntos de variáveis na matriz geográfica: correlação canônica. In: FAISSOL, S. **Tendências atuais na geografia urbano/regional: teorização e quantificação**. Rio de Janeiro: IBGE, 1978. p. 133-144
- ALBUQUERQUE, V. **Medidas de Posição Central**. Notas de Aula. PUC, RG. Disponível em: <http://www.pucrs.br/famat/valter/Administracao/4_Medidas_de_Posicao.doc>. Acessado em 27 de dezembro de 2007.
- ALVES, H.; CUNHA, M. L. **Uma introdução a alguns aplicativos, numa abordagem inicial dos dados**. Software Estatístico. Dossiê Didático. Disponível em: <<http://alea-esp.ine.pt>>. Acesso em 22 de Dezembro de 2007.
- ANJOS, A. **Planejamento e Experimentos II**. Curitiba, Universidade Federal do Paraná. 2004.
- BARBIN, D. **Componentes de variância: teoria e aplicações**. Piracicaba: FEALQ, 2004. 120p.
- BENJAMINI, Y. Opening the Box of a Boxplot. **The American Statistician**, Alexandria v. 42, n. 4, p. 257-262, 1988.
- BEST, D. J.; RAYNER, J. C. W. Smooth extensions of Pearsons's product moment correlation and Spearman's rho. **Statistics and Probability Letters**, Elsevier, v. 30, n. 2, p. 171-177, 1996.
- BOCK, R. D. **Multivariate statistical methods in behavioral research**. New York: McGraw-Hill, 1975. 623p.
- BUSSAB, W. O.; MORETTIN, P. A. **Estatística Básica**. São Paulo: Atual, 1987. 321 p.

CAMPOS, G. M.; BARBIN, L. E.; SPANÓ, E. C. J.; SILVA, S. R.; PÉCORÁ, D. J. **Estatística Prática para Docentes e Pós-Graduandos**: Tipos de variáveis. Disponível em: <http://www.forp.usp.br/restauradora/gmc/gmc_livro/gmc_livro_cap_02.html>. Acesso em 23 de setembro de 2007.

CARVALHO, R. D. **Árvore de decisão / algoritmo genético para tratar o problema de pequenos disjuntos em classificação de dados**. Rio de Janeiro, 2005, 173f. Tese (Doutorado em Ciências Engenharia civil). Universidade Federal do Rio de Janeiro – UFRJ.

CARVALHO, D. R.; FREITAS, A. A.; EBECKEN, N. F. F. **A critical review of rule surprisingness measures**. Rio de Janeiro, 2003. Disponível em http://www.cs.kent.ac.uk/people/staff/aaf/pub_papers.dir/ICDM-Rio-2003.pdf>. Acesso em : 23 de setembro de 2007.

CARVALHO, R. B. **Árvore de Decisão / Algoritmo Genético para tratar o problema de pequenos disjuntos em classificação de dados**. Rio de Janeiro, 1999, 173 f. Tese, (Doutorado) - Universidade Federal do Rio de Janeiro, UFRJ.

CAVADIAS, G.; PILON, P.; YUE, S. Power of the Mann-Kendall and Spearman's rho tests for monotonic trends in hidrological series. **Journal of Hydrology**, v. 259, p. 254 -271, 2002.

CHOONPRADUB, C.; MCNEIL, D. Can the box plot be improved? **Songklanakarin Journal of Science and Technology**, Hat Yai, v. 27, n. 3, p. 649-657, 2005.

COELHO, P. S. S. **Um sistema para indução de modelos de predição baseados em árvores**. Rio de Janeiro, 2005, (Tese de Doutorado). Universidade Federal do Rio de Janeiro, UFRJ.

COCHRAN, W. G.; COX, G. M. **Experimental designs**. 2 ed. New York: John Wiley & Sons, 1978. 611 p.

COSTA, S. C. **Modelos lineares generalizados mistos para dados longitudinais**. Piracicaba, 2003,(Tese de Doutorado) – ESALQ, Universidade de São Paulo.

COSTA, N. P. O. **Estatística**. São Paulo: Edgard Blucher, 1977. 266 p.

DEMETRIO, C. G. B. **Modelos lineares Generalizados em experimentação agrônômica**. Piracicaba: ESALQ, Departamento de Ciências Exatas, 2001. 113p.

DRAPER, N. R.; SMITH JR., H. **Applied Regression Analysis**. 3 ed. New York: John Wiley & Sons, 1998. 706 p.

EFRON, B; TIBSHIRANI, R. **Cross-Validation and the Bootstrap**: estimating the error rate of a prediction rule. Technical Report (Tr-477). Stanford: Stanford University, 1995.

ESTELBERGER, W; REIBNEGGER, G. The rank correlation coefficient: na additional AID in the interpretation of laboratory data. **Clinica Chimica Acta**, Graz, v. 239, n. 2, pag. 203-207, 1995.

FAYYAD, U. M. **Advances in Knowledge Discovery and Data Mining**. Menlo Park: AAAI Press/MIT Press, 1996. 611p.

FELDENS, M. A. **Descoberta de conhecimento aplicada à detecção de anomalias em base de dados**. Porto Alegre: UFRGS, 1999.

FERREIRA, A.; ALVES, L. **Inteligência Artificial e Data Mining**. Salvador, 2005, Dissertação (Mestrado) Universidade Federal da Bahia (UFBA).

GIOLO, S. R. **Análise de Regressão e Correlação**. Curitiba: Universidade Federal do Paraná. 2005.

GOMES, P. F. **Curso de estatística experimental**. 13 ed. Piracicaba: ESALQ-USP, 1990.

GONÇALVES, E. B.; CASTRO, I. M. Desenho esquemático para programas de qualidade em laboratórios de análises. **Ciência e Tecnologia de Alimentos**, Campinas, v. 18, n. 1, p. 121 – 126, 1998.

GUIMARÃES, P. B. **Análise Não-Paramétrica**. Curitiba: Universidade Federal do Paraná, 2003.

HAIR, J. F.; ANDERSON, R. E.; TATHAM, R. L.; BLACK, W. C. **Multivariate Data Analysis: With Readings**. 4 ed. Upper Sadle River: Prentice Hall, 1995.

HAN, J., KAMBER, M. **Data Mining: concepts and techniques** San Diego: Morgan Kaufmann, 2001. 500 p.

LIANG, K. Y.; ZEGER, S. L. Longitudinal data analysis using generalized linear models. **Biometrika**, Oxford, v.73, n.1, p.13–22. 1986.

MASSAD, E.; SILVEIRA, P. S. P. **Métodos quantitativos: conceitos iniciais**. Universidade de São Paulo (USP), 2003. Disponível em: <<http://www.dimfm.usp.br/metquant/index.php>> Acessado em: 22 de Dezembro de 2007.

MATTAR, F. N. **Pesquisa de marketing: metodologia, planejamento, execução e análise**. 2. ed. São Paulo: Atlas, 1994. 275 p.

MCCULLAGH, P.; NELDER, J. A. **Generalized linear models**. 2.ed. London: Chapman & Hall, 1989. 511 p.

MONTGOMERY, D. C. **Design and analysis of experiments**. 3ed. New York: John Wiley & Sons, 2001. 856 p.

MURARO, M. R. **Componentes físicos do sistema de raízes de pastagem de inverno formada pelo consócio aveia e azevém no sistema de integração lavoura pecuária**. Curitiba, 2004, (Dissertação de Mestrado) – UFPR, Universidade Federal do Paraná.

NELDER, J.; WEDDERBURN, R.W. Generalize linear models. **Journal Research Statistic Science**, v.135, p.370-384, 1972.

NETO, J. M. M. **Estatística multivariada: uma visão didática-metodológica**. Disponível em <http://criticanarede.com/cien_estatistica.html>. Acesso em 23 de fevereiro de 2007.

NETO, A. C. **Análise Multivariada** Curitiba: Universidade Federal do Paraná. 2005.

NETER, J.; KUTNER, M. H; NACHSTEIN, C. J.; WASSERMAN, W. **Applied linear statistical models**. Boston: Richard Irwin, 1996. 1408 p.

PEARSON, R. K. Exploring process data. **Journal of Process Control**, v. 11, n. 2, p. 179-194, 2001.

PÉREZ, L. F.; FREIRE, L. A. C. **Análise de dados longitudinais com aplicação na indústria.** (Relatório Técnico) 2003. Universidade Federal do Paraná. <<http://www.est.ufpr.br/rt/fl03.pdf>> Acesso em: 24 de Março de 2007.

PONTES, F. C. A. **Análise de variância multivariada com a utilização de métodos não-paramétricos e componentes principais baseados em matrizes de postos.** Piracicaba, 2005, 117pg. (Tese de Doutorado) - ESALQ, Universidade de São Paulo.

ROHATGI, V. K. **An Introduction to Probability Theory and Mathematical Statistic.** New York: Wiley, 1976. 684 p.

RUSSO, C. M. **Análise de um modelo de regressão com erros nas variáveis multivariado com intercepto nulo.** São Paulo, 2006. (Dissertação de Mestrado) – USP – São Carlos

RYAN, T.A., JOINER, B.L.. **Normal Probability Plots and Tests for Normality.** Pennsylvania: Pennsylvania State University, 1976. 13 p.

R DEVELOPMENT CORE TEAM. R. **Writing R extensions.** Disponível em< <http://cran.r-project.org/doc/manuals/R-exts.pdf>> Acesso em: Fevereiro de 2006.

R DEVELOPMENT CORE TEAM. R. **A language and environment for statistical Computing.** R Foundation for Statistical Computing, Vienna, Austria. Disponível em <http://www.R-project.org> >. Acesso em: Fevereiro de 2006.

SCHWERTMAN, C. N.; SILVA, R. **Identifying outleirs with sequential fences.** Computational Statistics & Data Analysis 51. p. 3800 – 3810, 2007.

SIEGEL, S. **Estatística não-paramétrica.** São Paulo: MacGraw-Hill, 1975. 350p.

SPIEGEL, M. R. **Estatística.** 3. ed. São Paulo: Makron Books, 1994. 639 p.

TOLEDO, G. L.; OVALLE, I. I. **Estatística Básica.** São Paulo: Editora Atlas, p.430-431, 1995.

WEARE, B. C., NASSTROM J. S. N. Examples of extended empirical orthogonal function analysis. **Monthly Weather Review**, Boston, v.110, n. 7, p. 481- 485, 1982.

WEKALIST. Interpret M5 model tree output. Disponível em:
<<https://list.scms.waikato.ac.nz/mailman/htdig/wekalist/2003-July/001609.html>>. Acesso em:
21 de fevereiro de 2007.

WEKA MACHINE LEARNING PROJECT. Disponível em:
<<http://www.cs.waikato.ac.nz/~ml>>. Acesso em: 10 de abril de 2007.

CAPÍTULO II - MODELOS LINEARES GENERALIZADOS, ÁRVORE DE CLASSIFICAÇÃO E ANOVA PARA ANÁLISE DE DADOS DE CONTAGEM DE ÁCARO RAJADO EM POMARES DE PESSEGUEIRO NA LAPA E CAMPO LARGO, PR..

MODELOS LINEARES GENERALIZADOS, ÁRVORE DE CLASSIFICAÇÃO E ANOVA PARA ANÁLISE DE DADOS DE CONTAGEM DE ÁCARO RAJADO EM POMARES DE PESSEGUEIRO NA LAPA E CAMPO LARGO, PR.

RESUMO

Ao realizar planejamento do experimento, é verificado que não há preocupação com os pressupostos dos métodos estatísticos aplicados, com a caracterização da variável mensurada e suas relações com outros fatores. Esta despreocupação faz com que as conclusões não sejam fidedignas aos dados. O objetivo deste trabalho foi identificar os fatores que influenciam a incidência de ácaro rajado em pomares de pessegueiros por meio de ANOVA, Modelos Lineares Generalizados para Dados Longitudinais e Árvore de Classificação. Os dados foram obtidos de experimentos realizados em pomares de pêssogo (*Prunus persica*) cv. Chimarrita sob sistema de Produção Integrada de Pêssego (PIP) durante o período de 2003 a 2006 nos municípios da Lapa e Campo Largo, Paraná. O delineamento foi em blocos completos casualizados com três repetições, em um arranjo de parcelas sub-subdivididas. No pomar da Lapa foram administradas cinco doses de nitrogênio (40, 80, 120, 160 e 200 kg.ha⁻¹ ano⁻¹ de N) e em Campo Largo quatro doses de nitrogênio (80, 120, 160 e 200 kg.ha⁻¹ ano⁻¹ de N) nas sub-parcelas e duas doses de potássio (55 e 110 kg.ha⁻¹ ano⁻¹ de K₂O) nas sub-subparcelas. Foram realizadas análises do teor foliar (macronutrientes N, P, K, Ca e Mg g.kg⁻¹ e os micronutrientes Fe, Mn, Zn, Cu e B mg.kg⁻¹) para verificar a relação entre os macro e micronutrientes com a quantidade de ácaros. Foi aplicada a técnica de Árvore de Decisão para definir o grau de importância dos fatores, Modelos Lineares Generalizados para Dados Longitudinais (MLGDL) considerando distribuição de Poisson e ANOVA. O fator que apresentou maior relevância para definir o nível de ácaros rajados foi a localização das plantas, seguido pela adubação de nitrogênio no solo. MLGD e Árvore Classificação mostraram que o nível de ácaros é independente da adubação potássica e que à medida que se é aplicado nitrogênio (kg.ha⁻¹ ano⁻¹) no solo a tendência do nível de ácaro rajado é aumentar até 160 kg.ha⁻¹ ano⁻¹ não diferindo de 200 kg.ha⁻¹ ano⁻¹ de nitrogênio. O Modelo Linear Generalizado foi mais adequado para avaliar os dados de contagem de ácaros dentre as datas observadas, seguida pela Árvore de Classificação, pela sua simplicidade; a ANOVA não se mostrou adequada.

Palavras-chave: *Tetranychus urticae*, Nitrogênio, Potássio, MLG, Árvore de Classificação.

**GENERALIZED LINEAR MODELS, TREE CLASSIFICATION AND ANOVA
APPLIED ON LEVELS POPULATION OF SPIDER MITE IN ORCHARDS OF
PEACH IN LAPA END CAMPO LARGO, PR.**

ABSTRACT

By making planning of the experiment, there is concern about the assumptions of the statistical methods applied, as the characterization of the variable measured in its relations with other factors. This unconcerned it means that the conclusions of the data is not representative. Objective of this work is to identify the factors that influence the incidence of mites in orchards of peach through of analysis of variance, Generalized Linear Models in Longitudinal Data and Tree of Classification. The experiments were conducted in orchards of peach cultivar Chimarrita in Lapa and Campo Largo, PR – Brazil, both conducted in the system of Integrated Production of Peach (PIP) since 2003. The experiment was conducted in an arrangement of split-split plots in blocks complete randomized design with three replications, In the orchard of Lapa were administered five merger of nitrogen (40, 80, 120, 160 and 200 kg.ha⁻¹ years⁻¹ of N), in Campo Largo four merger of nitrogen (80, 120, 160 and 200 kg.ha⁻¹ years⁻¹ of N), in the sub-plots and two concentrations of potassium (55 and 110 kg.ha⁻¹ years⁻¹ of K₂O) in the sub-subplots. We performed analyses of soil and leaf. Applied was the technique Tree of Classification to define the degree of importance of factors, Generalized Linear Models in Longitudinal Data (MLGLD) considering the Poisson distribution and analysis of variance. The factor that showed greater relevance to set the level of spider mites was the location of the plant, after by fertilizer nitrogen in the soil. MLGLD and Tree of Classification showed that the level of mites is independent of fertilization and that as it is applied nitrogen (kg.ha⁻¹ years⁻¹) on the ground trend in the level of spider mite is to increase 160 kg.ha⁻¹ years⁻¹ 200 kg.ha⁻¹ years⁻¹ not differing from of nitrogen. Analysis of Variance was not adequate to assess the level population of mites among the dates observed. MLGLG was more appropriate for decision after by the Tree of Classification for its simplicity.

Key word: *Tetranychus urticae* Nitrogen, Potassium, Tree of Classification

1 INTRODUÇÃO

As ferramentas estatísticas para análises dos dados são de uso corrente para a análise e interpretação de resultados. Contudo, observa-se que estas não são utilizadas adequadamente, sendo um dos possíveis motivos a diversidade de técnicas estatísticas disponíveis atualmente (COSTA, 2003).

A velocidade com que se desenvolvem as inovações em informática favorece o crescimento de *softwares* e *hardwares*, entre outros equipamentos com elevada tecnologia (CARVALHO, 2005), os quais são eficientes em coletar, organizar e armazenar grandes quantidades de dados de operações diárias e/ou de pesquisas científicas. Apesar disso, as organizações públicas e privadas brasileiras tiveram intensificaram acesso a estas tecnologias na década de 90, o que prejudicou a difusão do conhecimento (CARVALHO, FREITAS e EBECKEN, 2003).

Com o desenvolvimento da informática, uma das ciências que mais se beneficiou foi a Estatística. A aplicação prática de diversos métodos de análises estatísticas, de cálculos complexos permitiram a escolha de modelos matemáticos que melhor representem o conjunto de dados (HARRISON, 1998). Em função da diversidade de problemas e combinações analíticas, há dificuldade na escolha das ferramentas estatísticas para transformar dados em conhecimentos e subsidiar decisões comerciais e/ou científicas (FELDENS, 1999; FERREIRA e ALVES, 2005; ALVES e FERREIRA, 2006;). Neste ponto, verifica-se que na Entomologia e Ciências Agrárias não requerem cálculos estatísticos complexos e *softwares* sofisticados, porém, em muitas vezes não se verifica todos os pressupostos necessários para a validação das análises.

Em experimentos agrônômicos, na maioria das vezes, o interesse resume-se em verificar a influência dos fatores controláveis na variável resposta. Nota-se que não há preocupação com heterogeneidade da parcela experimental apresentada durante a implantação dos experimentos, muitas vezes sendo insuficiente a técnica de blocagem para evitar os efeitos não controláveis sobre a variável aleatória. O mesmo ocorre na caracterização da variável e suas relações com fatores, até então desconhecidos, que possivelmente possam a vir influenciar na resposta.

Nessas condições, as conclusões podem não representar fidedignamente os dados obtidos a campo. Por exemplo, ao estudar a evolução de ácaros durante um período determinado, a contagem destes artrópodes seria a variável resposta, assim cada árvore

monitorada representaria a unidade amostral e a condição de avaliação seria o tempo. Em outros estudos, a avaliação de ácaros situados a diferentes pomares teria como variável resposta o próprio nível populacional, e os locais representariam as unidades amostrais e a condição de avaliação seria a distância entre eles. Em ambas as situações, pode se considerar que a contagem de ácaros segue uma distribuição de Poisson, devido ao excesso de resposta zero e a própria característica da variável como proposto no trabalho de CINCONTI et al., 2005.

Para evidenciar a importância da caracterização da variável resposta, o presente trabalho utilizou dados obtidos em dois experimentos que estudaram a flutuação populacional de ácaros submetido a diferentes doses de adubação nitrogenada e potássica. O objetivo foi analisar os dados de contagem de ácaros, comparando os resultados dos três métodos estatísticos: Análise de Variância, Modelos Lineares Generalizados para Dados Longitudinais e Árvore de Classificação.

2 MATERIAL E MÉTODOS

2.1 MATERIAL

Os dados utilizados neste trabalho foram provenientes de dois experimentos em pomares comerciais de pessegueiro (*Prunus persica*) cv. Chimarrita, submetidos a diferentes doses de adubação nitrogenada e potássica, realizados por alunos doutorandos do Programa de Pós-Graduação em Agronomia, área de concentração em Produção Vegetal e componentes do Grupo de Ensino, Extensão e Pesquisa em Produção Integrada – GEEPPi da Universidade Federal do Paraná.

Os ensaios foram conduzidos nas safras de 2003/04 e 2004/05 em Campo Largo, PR e 2003/04, 2004/05 e 2005/06 na Lapa, PR.

O pomar na Lapa localiza-se nas coordenadas: latitude 25°45'35" S, longitude 49°43'47" W e altitude de 908 metros e o clima classificado (Köppen) como Cfb (subtropical úmido), com ocorrência de geada. Na figura 1 é apresentado as temperatura (°C) máximas, mínimas e média entre os anos de 2003 a 2006.

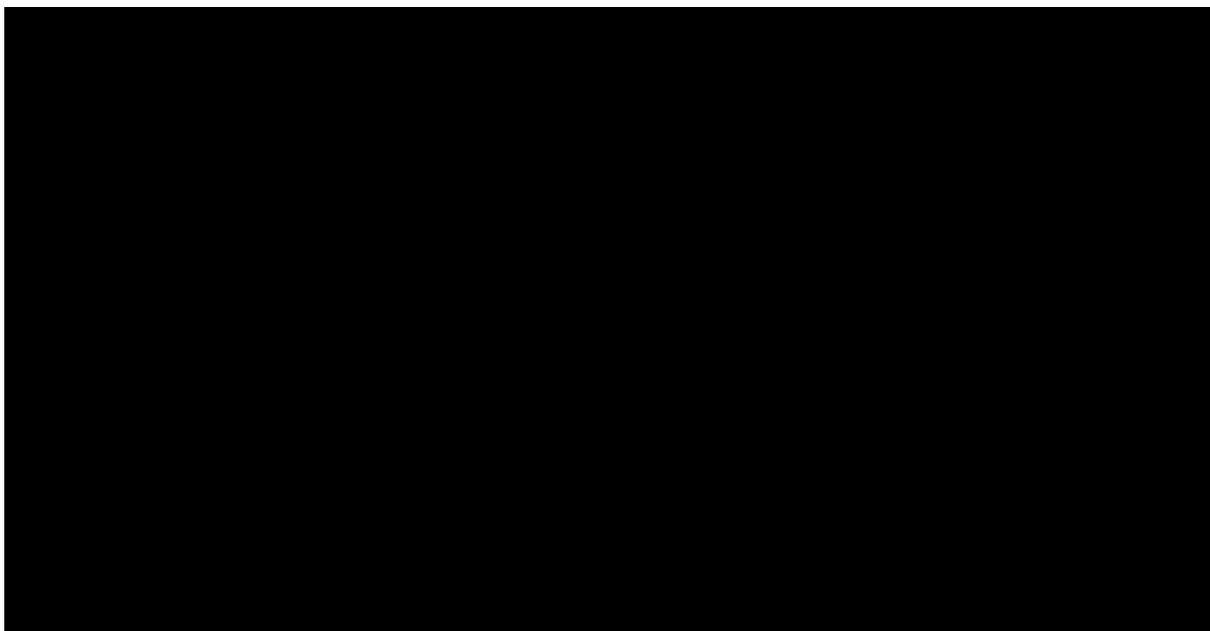


FIGURA 1 - DADOS DE TEMPERATURA MÁXIMA, MÍNIMA E MÉDIA DIÁRIA (°C), DA ESTAÇÃO METEOROLÓGICA SIMEPAR MAIS PRÓXIMA DA ÁREA EXPERIMENTAL

FONTE: SIMEPAR - Sistema Meteorológico do Paraná

O pomar é constituído de solo Cambissolo + Neossolo, com afloramentos rochosos e relevo ondulado. Para plantio de pêsego o solo foi escarificado e arado, sendo que a

adubação de base com fosfato e homogeneizada com gradagem. A análise do estado nutricional do solo foi realizada por meio de amostragens aleatórias situadas a 0 – 20 cm de profundidade, 20 cm do tronco (PT) e um terço da projeção da copa (PC), em cada parcela. As amostras, para análise química de rotina, foram coletadas em 14/07/2004 e encaminhadas a Fundação ABC, Castro (Pr) (Tabela 1A). A variedade Chimarrita foi plantada no espaçamento 5 x 2,5m e a condução foi em taça.

TABELA 1A - ANÁLISE QUÍMICA DO SOLO EM POMAR DE PESSEGUEIRO NO PERIMETRO DO TRONCO (PT), NA PROJEÇÃO DA COPA (PC), A 0 - 20 CM DE PROFUNDIDADE, EM 2004 (LAPA - PR).

Local	Prof. ¹	P resina	M.O.	pH	H+Al ⁺³	Al ⁺³	K ⁺	Ca ⁺²	Mg	SB	CTC	V
		mg/dm ³	g/dm ³	CaCl ₂	-----mmolc/dm ³ -----							
PC	00 - 20	19,9	31,5	6,6	15,3	0,0	3,4	97,9	55,5	156,7	172,0	90,5
PT	00 - 20	38,3	31,6	6,6	16,8	0,0	3,7	99,2	43,5	146,4	163,2	89,4
Média		29,1	31,6	6,6	16,1	0,00	3,5	98,5	49,5	151,6	167,6	89,9

NOTAS: ¹Prof. Profundidade da coleta em cm. M.O. Matéria orgânica

O pomar de Campo Largo, situa-se na latitude 25°29'75" S, longitude 49°35'21" W e altitude de 925 metros e o clima classificado (Köppen) como Cfb (subtropical úmido), com ocorrência de geada. O pomar é constituído de solo Cambissolo; o pomar foi escarificado e arado durante a instalação, simultaneamente a adubação de base foi fosfato. A análise de solo, nas profundidades de 0 - 20 cm e 20 – 40 cm foi, realizada na implantação do experimento. As amostras, para análise química de rotina, foram encaminhadas a Fundação ABC, Castro (Pr) (Tabela 2). A variedade Chimarrita foi plantada no espaçamento 4,0 x 1,25m e condução foi em Y.

TABELA 2 - ANÁLISE QUÍMICA DO SOLO EM POMAR DE PESSEGUEIRO NO PERIMETRO DO TRONCO (PT), NA PROJEÇÃO DA COPA (PC) E NA ENTRE LINHA (EL), A 0 - 20 E 20 - 40 CM DE PROFUNDIDADE, EM 2003 (CAMPO LARGO - PR).

Local	Prof. ¹	pH	Al ⁺³	H+Al ⁺³	Ca ⁺² +Mg ⁺²	Ca ⁺²	K ⁺	T	P	C	pH	V
		CaCl ₂	-----cmol _c /dm ³ -----								mg/dm ³	g/dm ³
PT	00 – 20	4,9	0,2	5,8	4,39	3,4	0,4	10,6	21,3	27,5	5,80	45,44
PT	20 – 40	4,7	0,5	6,7	3,56	2,5	0,1	10,4	7,9	20,8	5,60	35,33
PC	00 – 20	5,3	0,0	4,6	6,33	4,6	0,4	11,3	9,1	26,9	6,10	59,29
PC	20 – 40	5,0	0,0	5,4	4,10	2,7	0,1	9,63	3,7	23,2	5,90	43,93
EL	00 – 20	5,1	0,0	5,0	5,05	3,6	0,3	10,3	6,1	27,5	6,00	51,60
EL	20 – 40	5,1	0,0	5,0	4,16	2,6	0,1	9,3	4,2	19,6	6,00	46,24
Média		5,0	0,1	5,4	4,60	3,2	0,2	10,3	8,72	24,25	5,90	46,97

NOTAS: ¹Prof. Profundidade da coleta em cm.

Os delineamentos experimentais foram em três blocos ao acaso. Os tratamentos se constituíram de diferentes doses de nitrogênio (N) e potássio (K). Na Lapa, os tratamentos são

formados por cinco doses de TN (40, 80, 120, 160 e 200 kg.ha⁻¹) e duas de TK (55 e 110 kg.ha⁻¹). Em Campo Largo, o experimento foi conduzido com quatro doses de TN (80, 120, 160 e 200 kg.ha⁻¹) e duas de TK (55 e 110 kg.ha⁻¹). A Figura 2 mostra o delineamento experimental dividido em sub-parcelas e sub-subparcelas, sendo que a parcela principal foi considerada as datas de contagem de ácaros, nos três anos de pesquisa.

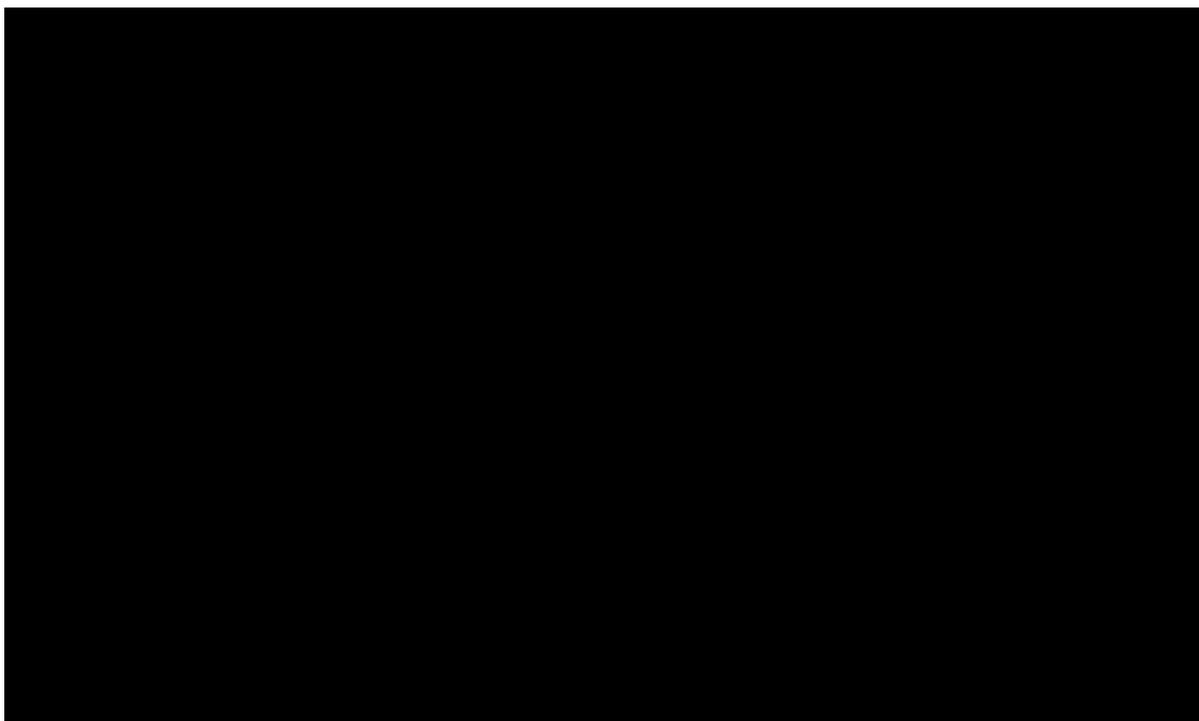


FIGURA 2 - DISPOSIÇÃO DE PÊSSEGO EM SUBPARCELAS E SUB-SUBPARCELAS CONSTITUÍDA DE DOSES DE NITROGÊNIO E POTÁSSIO EM DOIS POMARES, SITUADOS NA LAPA E CAMPO LARGO (PR), 2003 a 2006.

O nitrogênio na forma de uréia, foi aplicado com um dosador, previamente calibrado, na projeção da copa (FACHINELLO et al., 2003), sendo que 30% da concentração foi colocada na fase fenológica do início da brotação do pessegueiro, 30% no raleio e 40% após a colheita. O potássio foi aplicado na forma de cloreto de potássio, depositado nas mesmas condições para N, no início da brotação, na razão de 60%, e o restante após o raleio.

A avaliação nutricional dos pessegueiros foi feita por meio de teores foliares (macronutrientes N, P, K, Ca e Mg g.kg⁻¹ e os micronutrientes Fe, Mn, Zn, Cu e B mg.kg⁻¹). Em cada sub-subparcela foram coletadas 80 folhas completas (limbo + pecíolo), localizadas na parte mediana dos ramos do ano e a 1,7 m de altura e em todo perímetro, na

décima quarta semana após plena floração, nas safras de 2003/04 e 2004/05 em Campo Largo e 2003/04, 2004/05 e 2005/06 na Lapa. As folhas foram secas em estufa com circulação forçada de ar a $60 \pm 2^\circ\text{C}$ até massa constante, e as amostras foram encaminhadas para o Laboratório de Nutrição Mineral de Plantas, da Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA Clima temperado), Pelotas, RS e analisada segundo metodologia descrita por SILVA (1999).

O desenvolvimento de ácaros foi avaliado por meio da coleta de 10 folhas por planta, situadas em ramos do ano e a cerca de 1,7 m de altura. Estas foram acondicionadas em sacos de papel e transportadas ao laboratório em caixas de isopor com gelo. A contagem dos espécimes foi realizada com auxílio de estereomicroscópio no Laboratório em Manejo Integrado de Pragas da UFPR. As coletas na Lapa ocorreram em 20/01/2004, 17/02/2004, 12/12/2004, 02/01/2005, 21/01/2005, 09/02/2005, 16/12/2005, 05/01/2005, 23/01/2006, 11/02/2006, 08/03/2006, e em Campo Largo nos dias 28/12/2003, 23/1/2004, 12/12/2004, 2/1/2005, 21/1/2005, 9/2/2005.

2.2 MÉTODOS

Os dados de contagem de ácaros nas diferentes datas foram submetidos a Análises de Variância, Árvore decisão e Modelos Lineares generalizados para dados longitudinais.

Para a aplicação de análise de variância se pressupõe que haja independência entre as variáveis respostas e, quando necessário, foi realizada a transformação de Box-Cox (JUNIOR, 2005) para atender os pressupostos dessa análise (MONTGOMERY, 2001). O Teste Tukey ($p\text{-valor} < 0,05$) foi utilizado para Teste F significativo (BARBIN, 2004).

Para verificar os pressupostos da ANOVA foi realizado análise gráfica de saída padrão do *software* R e os métodos de Shapiro-Wilk para normalidade, Bartlett para homocedasticidade dos resíduos.

Para viabilizar o uso da Análise Árvore de Classificação na Lapa, realizou-se o *Bootstrap* (EFRON e TIBSHIRANI, 1995) na amostra realizada em 16/12/2005, a qual gerou mais 15 observações. Com base no número médio de ácaros móveis, foi realizada uma discretização dos valores para aplicação da metodologia de Árvore de Classificação (WITTEN e FRANK, 2000), do qual se definiu as classes de 0 a 43, 44 a 100 e acima de 100 ácaros, respectivamente assumindo os níveis 1, 2 e 3. Para ajustar o modelo, realizou-se seis treinos a partir de seis sub-amostras, contendo 44 observações e seis sub-amostras contendo uma observação para teste. Em Campo Largo utilizou-se toda a base de informações,

formando as classes 0, 1 a 2 e acima de 3 ácaros, respectivamente assumindo os níveis 1, 2 e 3.

Para ajustar o modelo realizou-se seis treinos a partir de seis sub-amostras, contendo 143 observações e seis sub-amostras contendo uma observação para teste.

A teoria de modelos lineares generalizados (MCCULLAGH e NELDER,1989; DEMETRIO, 2001) foi utilizada para explicar a influência dos tratamentos sobre o desenvolvimento de ácaros, considerando a função da Distribuição Probabilística de Poisson.

Foi utilizada a Análise de Correlação de Pearson para verificar dependência entre as populações de ácaros, precipitação e temperatura e a Análise de Correlação de Spearman para relacionar os tratamentos com população de ácaros (SIEGEL, 1975; SPIEGEL, 1994; GIOLO, 2005).

As análises estatísticas aplicadas aos dados foram realizadas pelo *software* WEKA (*Waikato Environment for Knowledge Analysis*), *software* R (*R Development Core Team* 2005) e os *softwares* MSTAT, SPSS.

3 RESULTADOS E DISCUSSÃO

Os resultados serão apresentados em uma seqüência de análises de modo que permita a comparação dos resultados obtidos pelos diferentes métodos aplicados a uma mesma variável. Iniciou-se por uma análise descritiva dos dados do pomar de pessegueiro, enfatizando as principais estatísticas. Em seguida, realizou-se análise de dispersão dos dados por meio do *Box-plot*, abordando individualmente cada safra e posteriormente em conjunto. Realizou-se Análise de Variância (ANOVA) e as conclusões são apresentadas independentemente dos pressupostos. Finalmente são apresentados Análise de Modelos Lineares Generalizados e Árvore de Classificação.

3.1 ANÁLISE DESCRITIVA

Lapa. Na safra de 2003/04, a população de ácaros sobre as folhas foi baixa, com média inferior 0,4 ácaros por folha, sendo que o coeficiente de variação (C.V) foi elevado, independente dos tratamentos (Tabela 1). Em 2004/05, houve uma alta freqüência de valores zero, entretanto, a média de ácaros foi superior apresentada em 2003/04, com CV semelhante à safra anterior. Nesse ciclo vegetativo observou-se que a maior média de ácaros por folha ocorreu na concentração de $80 \text{ kg.ha}^{-1} \text{ ano}^{-1}$ nitrogênio (TN 80). Na safra de 2005/06, a população de ácaros cresceu acentuadamente, entretanto, a freqüência (Moda) permaneceu com o mesmo comportamento, nesse caso o maior número médio de ácaros ocorreu na concentração de $200 \text{ kg.ha}^{-1} \text{ ano}^{-1}$ de N (TN 200).

Observa-se que houve crescente desenvolvimento de ácaros ao longo dos três anos e somente no último ano se obteve uma tendência de desenvolvimento de ácaros em relação à concentração de N. A variável resposta se caracteriza, no pomar da Lapa, pela freqüência de zeros e dispersão dos dados em torno da média (Tabela 1).

Campo Largo. Na safra de 2003/04 e 2004/04, a população ácaros sobre as folhas foi baixa, com média inferior 2 ácaros por folha, entretanto os CV's foram elevados, semelhantes entre as safras, independente das doses de adubações. Não ocorreu aumento de ácaros do primeiro para o segundo ano. Obteve-se mesma conclusão do pomar na Lapa, alta freqüência de valores iguais a zeros e a alta dispersão dos dados em torno da média (Tabela 1).

TABELA 3 - DESCRITIVO DO NÍVEL POPULACIONAL DE ÁCAROS MÓVEIS SOB AS DIFERENTES DOSES DE ADUBAÇÕES NITROGENADAS E POTÁSSICAS ($\text{kg}\cdot\text{ha}^{-1}\text{ ano}^{-1}$), E BLOCOS, REALIZADAS NAS SAFRAS DE 2003/04, 2004/05 E 2005/06 NO POMAR DE PÊSSEGO, LAPA E CAMPO LARGO, PR.

Trat $\text{kg}\cdot\text{ha}^{-1}\text{ ano}^{-1}$	Safrá 2003/04						Safrá 2004/05						Safrá 2005/06					
	Nº	Me	Md	Mo	Max	CV (%)	Nº	Me	Md	Mo	Max	CV (%)	Nº	Me	Md	Mo	Max	CV (%)
Pomar Lapa																		
Nitrogênio																		
8	230	30	12,300	0,000	0,000	102	227	40	12	0,028	0,000	0,000	0,3	272	24	0,22	0,00	0,00
24	275	30	18,800	0,600	0,000	163	217	80	12	0,083	0,000	0,000	0,3	144	24	21,63	0,00	0,00
28	233	30	25,000	1,250	0,000	245	227	120	12	0,427	0,000	0,000	3,8	253	24	3,06	0,00	0,00
48	177	30	20,000	1,750	0,000	192	186	160	12	0,005	0,000	0,000	0,1	346	24	6,81	0,25	0,00
33	333	30	36,817	0,750	0,000	324	214	200	12	0,036	0,000	0,000	0,2	200	24	2,19	0,00	0,00
Potássio																		
24	367	75	25,260	1,000	0,000	324	224	55	30	0,183	0,000	0,000	3,8	380	60	8,45	0,00	0,00
90	465	75	23,747	1,000	0,000	246	223	110	30	0,048	0,000	0,000	0,3	187	60	5,39	0,00	0,00
Bloco																		
11	333	50	11,480	0,000	0,000	90	201	1	20	0,009	0,000	0,000	0,2	447	40	0,58	0,00	0,00
24	367	50	28,500	0,500	0,000	245	195	2	20	0,234	0,000	0,000	3,8	355	40	12,41	0,00	0,00
78	206	50	33,530	1,250	0,000	324	214	3	20	0,103	0,000	0,000	0,9	213	40	7,78	0,00	0,00
24	406	150	24,503	1,000	0,000	324	223	Total	60	0,116	0,000	0,000	3,8	430	120	6,92	0,00	0,00
Pomar Campo Largo																		
Nitrogênio																		
3	163							80	12	1,667	0,500	0,000	6,0	131	24	0,54	0,00	0,00
19	247							120	12	0,583	0,000	0,000	2,0	154	24	1,58	0,00	0,00
12	130							160	12	1,000	0,000	0,000	4,0	148	24	2,92	1,50	0,00
7	90							200	12	1,917	1,000	0,000	13,0	192	24	2,21	2,00	0,00
Potássio																		
12	171							55	24	1,042	0,000	0,000	5,0	143	48	1,42	0,00	0,00
19	158							110	24	1,542	0,000	0,000	13,0	188	48	2,21	1,00	0,00
Bloco																		
12	163							1	16	2,125	1,000	0,000	13,0	160	32	2,00	1,00	0,00
19	156							2	16	1,188	0,000	0,000	5,0	148	32	2,22	2,00	0,00
9	178							3	16	0,563	0,000	0,000	2,0	129	32	1,22	0,00	0,00
19	166							Total	48	1,292	0,000	0,000	13	178	96	1,81	1,00	0,00

Trat. = doses de adubação; Nº = Número de observações; Me = Média; Md = Mediana; Mo = Moda; Max = Valor máximo; CV (%) = Coeficiente de Variação

Verifica-se que análise descritiva, realizada nos dois pomares, facilita a interpretação dos resultados, descrevendo-os resumidamente com suas principais estatísticas (TOLEDO e OVALLE, 1995). Podem identificar características que possam evidenciar futuros problemas na validação dos pressupostos quando da utilização da ANOVA. Neste caso, o uso de outra metodologia de análise estatística que não seja fundamentada pela distribuição normal e/ou pela homocedasticidade dos resíduos.

Complementando a análise descritiva, a qualidade dos dados pode ser verificada pela dispersão gráfica, com utilização do *Box-plot* (BUSSAB e MORETTIN, 2002). O qual

As dispersões apresentadas em ambos os pomares, poderão influenciar os pressupostos exigidos pela ANOVA, o que obriga ao pesquisador a suspeitar que a variabilidade influencie as análises. A utilização da ANOVA nestas condições levará a um número elevado de contrastes não significativos. Logo, para definição do método de análises estatísticas deve ser considerado o seguinte: incidências de zeros, variável resposta discreta e finita, variabilidade dos dados e fatores (data, bloco e tratamentos) (COCHRAN e COX, 1978; BANZATTO e KRONKA, 1989; GOMES, 1990). Esses autores sugerem minimizar o efeito do erro experimental por meio de uma melhor elaboração do delineamento visando melhorar a diferenças entre os tratamentos.

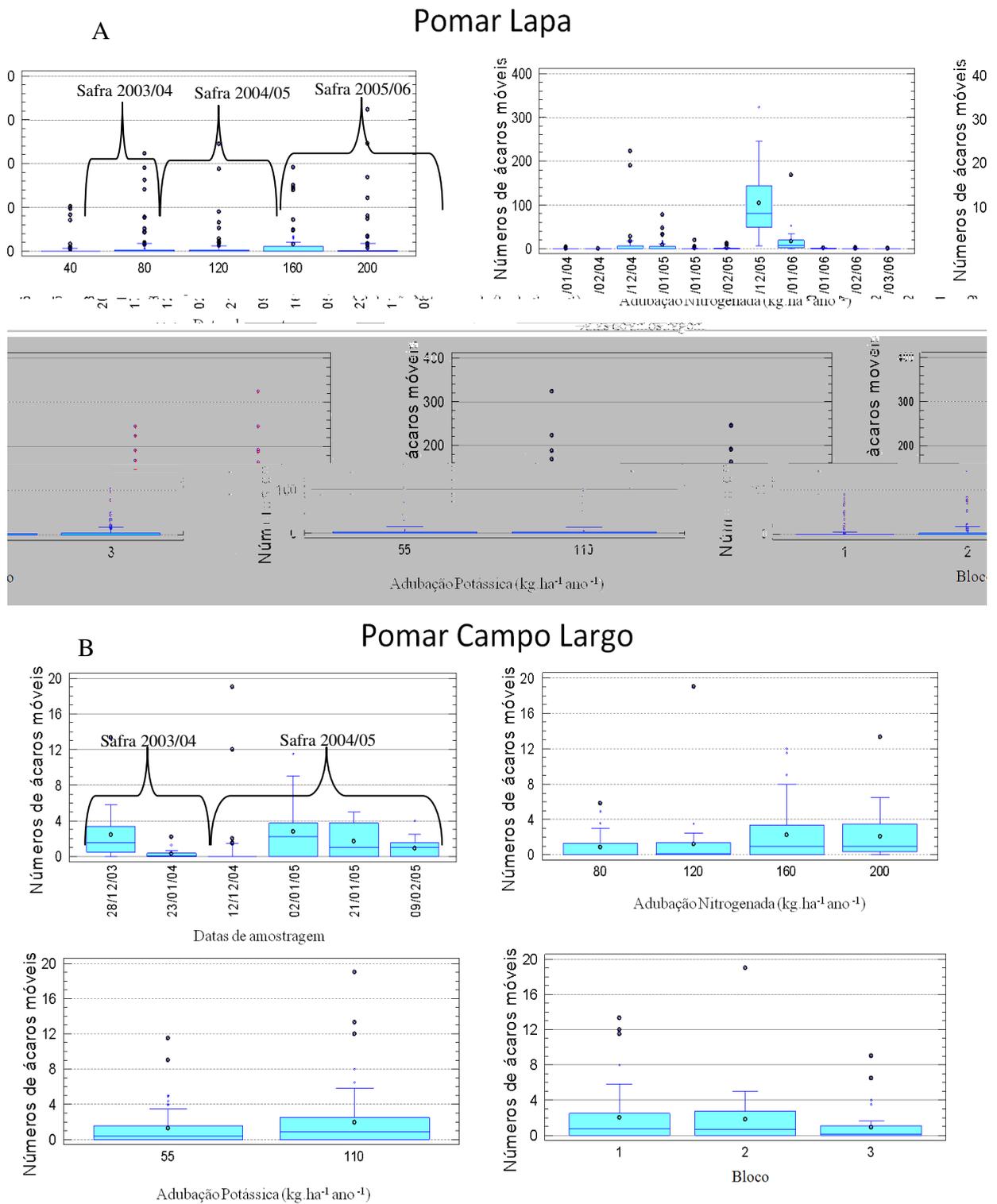


FIGURA 3 - *BOX-PLOT* DO NÍVEL POPULACIONAL DE ÁCAROS MÓVEIS SOB AS DIFERENTES DOSES DE ADUBAÇÕES NITROGENADAS E POTÁSSICAS ($\text{kg}\cdot\text{ha}^{-1}\cdot\text{ano}^{-1}$) REALIZADAS NAS SAFRAS DE 2003/04, 2004/05 E 2005/06 NO POMAR DE PÊSSEGO. LAPA, PR.

3.3 ANÁLISE DE VARIÂNCIA

Lapa. A contagem de ácaros submetidos à ANOVA mostra que o único fator significativo foi Ano (TABELA 2). Entretanto, somente a data 16/12/2005 diferenciou das demais (Tukey, p -valor $<0,05$), visualizado na Figura 2. Os demais fatores não apresentaram significância quando aplicado o Teste F.

Campo Largo. Não foi possível afirmar que há contrastes significativos entre os fatores testados pelo Teste F (TABELA 2).

TABELA 4 - ANÁLISE DE VARIÂNCIA (ANOVA) DO NÍVEL POPULACIONAL DE ÁCAROS MÓVEIS SOB AS DIFERENTES DOSES DE ADUBAÇÕES NITROGENADAS E POTÁSSICAS (kg ha^{-1} ano $^{-1}$), E BLOCOS, REALIZADAS NAS SAFRAS DE 2003/04, 2004/05 E 2005/06 NO POMAR DE PÊSSEGO DA LAPA E CAMPO LARGO, PR.

CAUSAS DE VARIÂÇÃO	DE	POMAR LAPA				POMAR CAMPO LARGO			
		GL	SQM	QM	P-valor	GL	SQM	QM	P-valor
Bloco		2	5620	2810	0,16	2	16,26	8,13	0,11
Ano (data)		10	28606	2860,6	0,01	5	20,15	5,037	0,43
Erro (bloco)		20	2780	139		10	5,71	0,571	
Parcela									
Adubação N		4	1378	344,5	0,26	3	16,50	5,5	0,69
Ano x Adubação N		40	1226	30,65	0,53	15	12,72	0,848	0,05
Erro (ano)		88	1032	11,727		36	6,42	0,1783	
Subparcela									
Adubação K		1	275	275	0,50	1	16,84	16,84	0,10
Ano x adubação K		10	88	8,8	0,98	5	6,87	1,374	0,33
Adub. N x adub. K		4	76	19	0,97	3	0,15	0,005	0,99
Ano x N x K		40	169	4,225	0,96	15	6,63	0,442	0,35
Erro (N)		110	597	5,427		48	5,82	0,12125	

GL = graus de liberdade, SQM= Somas dos Quadrados Médios, QM= Quadrado Médio, P-valor real do referente teste F.

* p -valor $< 0,05$ é significativo

Em ambos os pomares, os resultados não significativos podem ser explicados pela variabilidade elevada, verificado pelos quadrados médios dos resíduos (TABELA 2) (SIT, 1994), mostrando que as análises preliminares (TABELA 1 e FIGURA 2) são relevantes para indicar que a ANOVA pode ser ineficiente para rejeitar a hipótese nula, ou seja, que não haja diferenças significativas entre os níveis dos fatores.

A ANOVA deve ser utilizada para variáveis contínuas que seguem a distribuição normal (MASSAD E SILVEIRA, 2003). Para validar os resultados obtidos por meio ANOVA é necessário verificar os pressupostos, realizados por gráficos e métodos.

3.4 ANÁLISE DOS PRESSUPOSTOS DA ANOVA

Os gráficos gerados pelo *software* R são apresentados nas Figuras 3, 4, 5 e 6, que verificam os pressupostos de normalidade, homocedasticidade, independência do modelo gerado.

Lapa. Pela análise gráfica verifica-se que a variância dos resíduos é proporcional aos preditos (*Residuals vs Fitted values*) (Figura 3A).

A normalidade dos resíduos (*Standardized residuals vs Theoretical Quantiles*) não foi atendida, pois a curva é influenciada pelos pontos extremos (Figura 3B). O teste Shapiro Wilk (p-valor < 0,001) corrobora com a análise gráfica.

Os valores preditos perante aos desvios-padrão (*(Standardized)^{0,5} vs Fitted values*) tem a tendência de se concentrar no valor zero (Figura 3C).

O resultado obtido pelo teste de Bartlett com p-valor < 0,001, afirmando que não existe igualdade entre as variâncias foi confirmado pela análise gráfica (*Standardized residuals vs Factor Level Combinations*) (Figura 3D) (SNEDECOR et. al., 1989)

As discussões geradas pelas análises gráficas (Figura 3) e pelos métodos (Bartlett e Shapiro Wilk) identificaram discrepância nos dados de contagem de ácaros e para fundamentar as análises é necessário que os mesmos sejam transformados (JUNIOR, 2005).

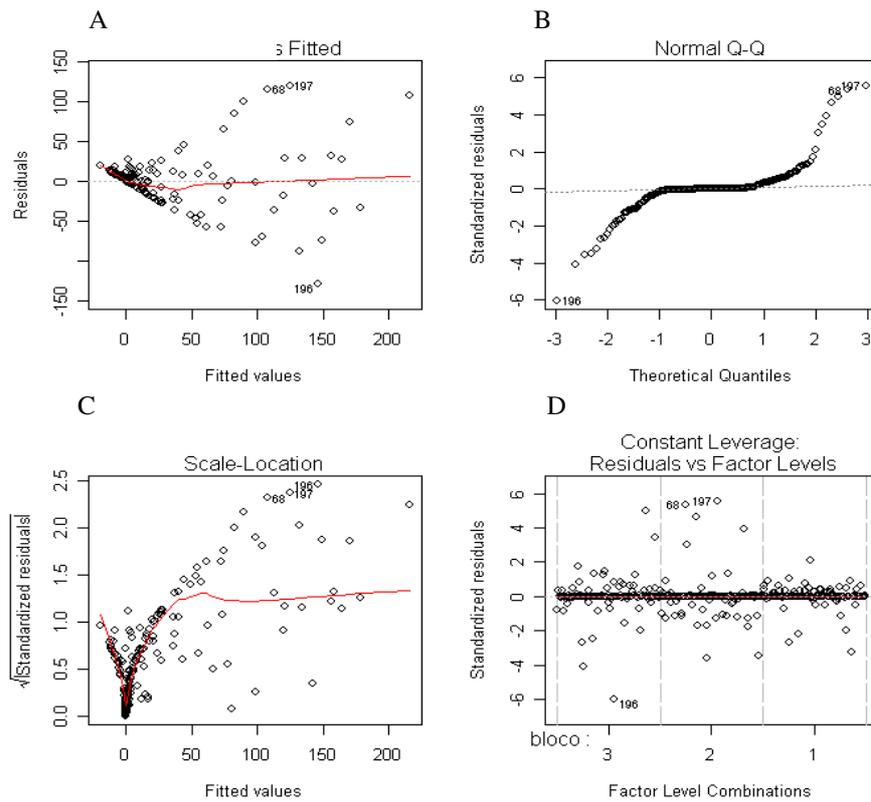
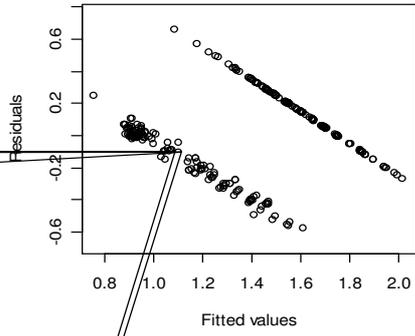
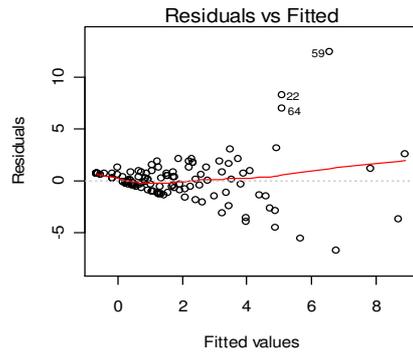


FIGURA 4 - ANÁLISE DOS PRESSUPOSTOS DA ANOVA SEM TRANSFORMAÇÃO DOS DADOS, LAPA, PR.

Os dados foram transformados pelo método de Box-Cox, conforme estudos realizados por AGUIRRE (1997), e o coeficiente de correção foi $\lambda = -0.02020202$.

A transformação dos dados não corrigiu a não-linearidade e heterocedasticidade. Os resultados de não-conformidade dos resíduos foram confirmados pelos métodos Shapiro Wilk e Bartlett, respectivamente com p-valores = 0.00461 e 0,0001 (Figura 4).





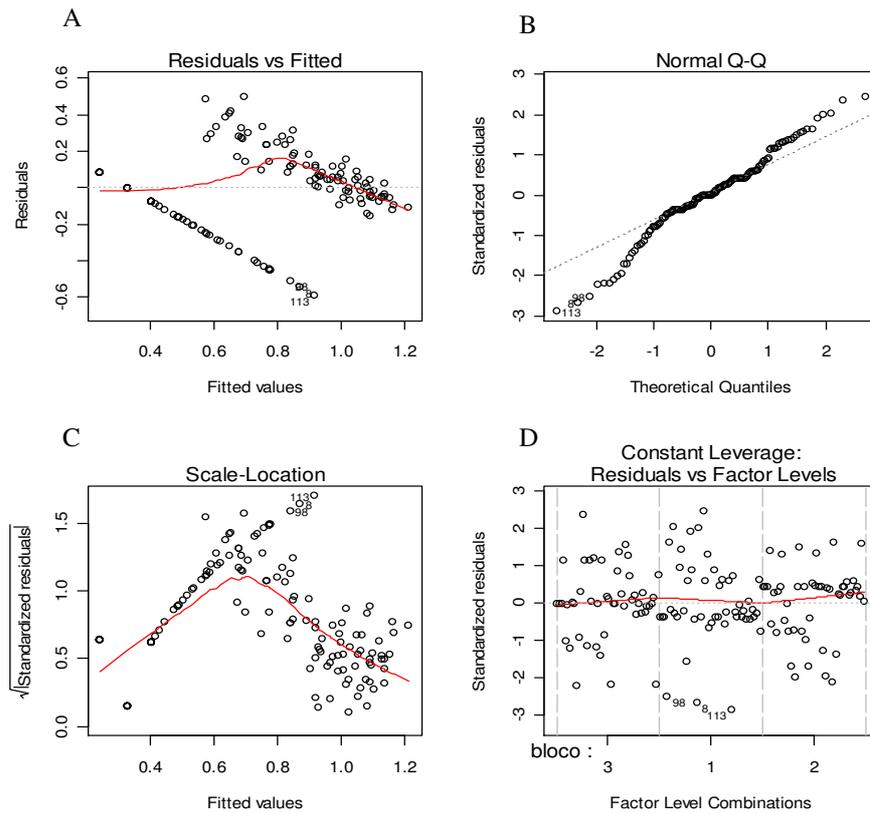


FIGURA 7 - ANÁLISE DOS PRESSUSPOSTOS DA ANOVA COM TRANSFORMAÇÃO BOX-COX $\lambda=0,06060606$, CAMPO LARGO.

Os dados de contagem de ácaros transformados de ambos os pomares, não deve ser analisados pela ANOVA, pois os resíduos não atenderam os pressupostos, podendo ser atribuído esse resultado à característica da variável que além de biológica é mensurada por contagem, bem como a outros fatores desconhecidos.

A ANOVA desenvolvida por FISHER requer verificar a validade de várias suposições fundamentais (BOWER, 2007). De acordo com HOGG e LEDOLTER (1987), para cada nível verificar se os valores seguem uma distribuição Normal e as variâncias são as mesmas. É comum analisar graficamente os pressupostos (RYAN e JOINER, 1976), porém usuários sem experiências têm dificuldades na interpretação dos gráficos. Portanto, métodos estatísticos como Bartlett e Shapiro Wilk, podem ser obtidos facilmente por *softwares* específicos em análises estatísticas (BOWER, 2007).

Outros estudos realizados pelo método de contagem de ácaros apresentam dificuldades na análise estatística, muitos dos quais não encontram diferenças entre os tratamentos sem

terem discutidos os pressupostos da ANOVA, ou não inseriram no texto. Como exemplo, a contagem de ácaros fitófagos em mamoeiro (*Carica papaya L.*) (VIERA et. al. 2004).

Para definir a metodologia de análise dos dados, é necessário identificar as características e o tipo da variável envolvida no experimento (CAMPO et. al., 2000), essas informações indicam os métodos estatísticos mais adequados. Dentre os diversos métodos, ANOVA é a mais utilizada em análises agrônômicas (ROHATGI, 1976), entretanto, é facilmente influenciada pela variabilidade acentuada e a não normalidade dos dados, levando a conclusões errôneas, visto que a ANOVA é eficiente para comparar dados contínuos (MONTGOMERY, 2001).

Experimentos que requeiram a mensuração de dados por contagem, na área agrícola, são analisados de forma mais adequada pelos modelos lineares generalizados (MCCULLAGH e NELDER, 1989; DEMETRIO, 2001).

3.5 MODELOS LINEARES GENERALIZADOS PARA DADOS LONGITUDINAIS

O modelo linear, fundamentado na Distribuição de Poisson, apresentado na Tabela 3 teve como objetivo avaliar os efeitos de uma ou mais variáveis explicativas/independentes sobre a variável resposta de interesse, ou seja, nível populacional de ácaros.

Lapa. Concluiu-se que o bloco III (p-valor =0,002) está ajudando a explicar o nível populacional de ácaros móveis no cultivo de pêssego na Lapa. Adubação nitrogenada e potássica não influenciaram no número de ácaros. Não foi possível identificar estrutura de correlação com as datas em que se realizou amostragem (Tabela 3).

Campo Largo. Os níveis de ácaros rajados está sendo explicado pelas adubações 160 TN, que apresentou a maior estimativa seguido pela adubação 200 TN. Com relação ao ácaro, vários estudos indicam que o nitrogênio aumenta a sua fecundidade, porém esse efeito varia consideravelmente dependendo da planta hospedeira (SUSKI e BADOWSKA, 1975). O bloco III está ajudando a explicar o nível populacional de ácaros móveis no cultivo de pêssego em Campo Largo. Também pode ser concluído que o nível de ácaros neste experimento não possibilitou identificar estrutura de correlação entre as datas em que se mensuraram os números de ácaros (TABELA 3), isso pode ser atribuído a intervenção aleatória realizadas nos pomares, bem como outros fatores desconhecidos.

TABELA 5 - ESTIMATIVA PARA O NÍVEL POPULACIONAL DE ÁCAROS MÓVEIS SOB AS DIFERENTES DOSES DE ADUBAÇÕES NITROGENADAS E POTÁSSICAS (K g ha⁻¹ ano⁻¹) REALIZADAS NAS SAFRAS DE 2003/04, 2004/05 E 2005/06 EM POMARES DE PÊSSEGO DA LAPA E CAMPO LARGO, PR.

POMAR LAPA				
β Parâmetro	Estimativa	Erro padrão	Estatística de Wald	P-valor
β ₀ Intercepto	1.946	0.504	14.88	0.001
β ₁ Bloco 2	0.032	0.369	0.007	0.932
β ₂ Bloco 3	1.203	0.392	9.383	0.001
β ₃ 80 kg.ha ⁻¹ Ano ⁻¹ de N	0.973	0.531	3.351	0.061
β ₄ 120 kg.ha ⁻¹ Ano ⁻¹ de N	0.705	0.566	1.550	0.210
β ₅ 160 kg.ha ⁻¹ Ano ⁻¹ de N	0.926	0.510	3.297	0.063
β ₆ 200 kg.ha ⁻¹ Ano ⁻¹ de N	1.039	0.565	3.379	0.069
β ₇ 110 kg.ha ⁻¹ Ano ⁻¹ de K	-0.13	0.330	0.163	0.682
β ₈ Datas	0.035	0.027	1.649	0.190
POMAR CAMPO LARGO				
β Parâmetro	Estimativa	Erro padrão	Estatística de Wald	P-valor
β ₀ Intercepto	-0.051	0.440	0.013	0.906
β ₁ Bloco 2	-0.109	0.319	0.117	0.731
β ₂ Bloco 3	-0.781	0.342	5.216	0.022
β ₃ 120 kg.ha ⁻¹ Ano ⁻¹ de N	0.329	0.511	0.414	0.519
β ₄ 160 kg.ha ⁻¹ Ano ⁻¹ de N	0.974	0.352	7.663	0.005
β ₅ 200 kg.ha ⁻¹ Ano ⁻¹ de N	0.884	0.337	6.861	0.008
β ₆ 110 kg.ha ⁻¹ Ano ⁻¹ de K	0.438	0.268	2.652	0.103
β ₇ Datas	-0.031	0.061	0.270	0.602

* p-valor < 0,05 é significativo

Segundo LIANG e ZEGGER (1986), citado por PÉREZ e FREIRE (2003) estudos longitudinais são conveniente quando o interesse é verificar variações globais ou individuais ao longo do tempo.

Observa-se na figura 3 que os teores foliares e os macros e micronutrientes do solo não ajudaram a explicar o nível populacional de ácaros em ambos os pomares.

Análise de modelos lineares generalizados foi mais eficiente que as análises de variância realizadas nos dados mensurados nos pomares da Lapa e Campo Largo, o qual identificou os fatores bloco e nitrogênio que ajudam a explicar a quantidade de ácaros móveis encontrados nas folhas de pessegueiro, corroborando com os resultados observados nas análises descritivas e dispersão.

Outra opção para análise do dados desse estudo é *Árvore de Decisão*, que não é fundamentada em distribuição probabilística, ou seja, é uma análise não paramétrica.

3.6 ÁRVORE DE CLASSIFICAÇÃO

Árvore de Classificação é uma análise não-paramétrica que apresentou os fatores hierarquicamente significativos, baseando-se em estágios de decisão e na separação de classes e subconjuntos (NOGUEIRA, 2007).

Lapa. O fator de maior relevância foi o bloco, ou seja, a localização das plantas definiu o nível populacional de ácaros. Portanto, é possível observar no pomar um agrupamento de ácaros. É possível afirmar que o terceiro bloco, independente das doses de adubação nitrogenada, definiu o nível populacional de ácaros. No bloco II o maior nível populacional de ácaros rajados foi encontrado na adubação de 160 TN e o menor nível em 120 TN, as demais adubações permaneceram no nível intermediário. A adubação de 200 TN obteve um nível de ácaros inferior à dose 160 TN devido à dispersão dos dados. No bloco III verifica-se que os maiores níveis de ácaros rajados ficaram entre adubações 120 TN e 160 TN e as menores nas doses entre 40 TN, 80 TN e 200 TN (Figura 7).

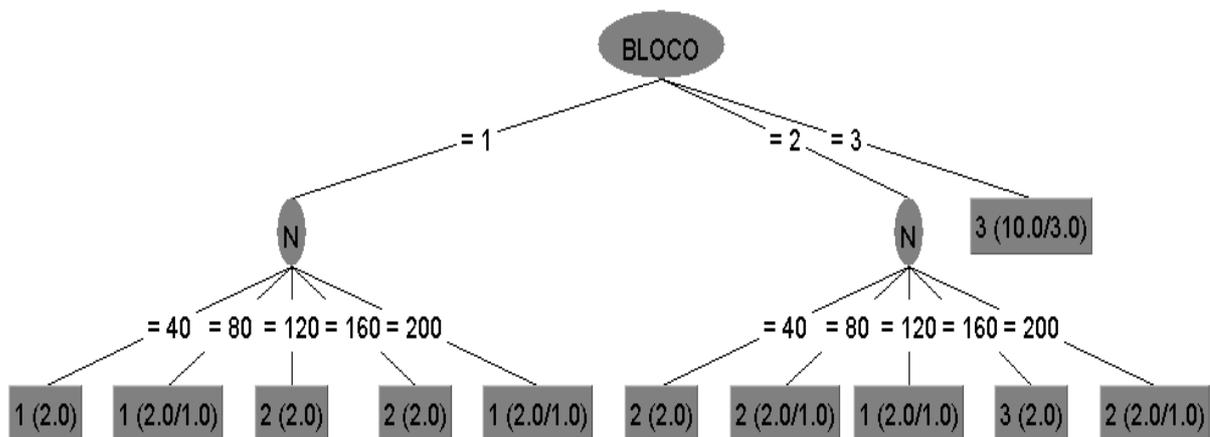


FIGURA 8 - ÁRVORE DE CLASSIFICAÇÃO COM 91,48% DE CONFIABILIDADE ORIGINADA DO TEOR FOLIAR E ADUBAÇÕES DE NITROGENIO E POTASSIO EM RELAÇÃO AO NÍVEL DE ÁCAROS RAJADOS AMOSTRADO NO POMAR DE PESSEGUEIRO LOCALIZADO NA LAPA, PR.

O comportamento dos ácaros em se concentrar no bloco III também foi observado nos resultados obtidos por OKU, YANO e TAKAFUJI (2002) em que a dispersão natural dos *T. urticae* esta relacionada positivamente com a indisponibilidade de alimentos, temperatura e

insetos, outro fator considerado pelo autor é características hereditárias da colônia (FADINI et al., 2007).

Segundo MITCHELL (1996) Árvore de Classificação é um método prático mais usado para implementar inferências indutivas, por meio de funções de saída discreta.

Campo Largo. O TN é principal fator que ajuda a responder a quantidade de ácaros móveis amostrados na folha. A adubação de 160 TN no bloco I e 200 TN no bloco II seguido pelo bloco III interagindo bloco II foram os que apresentaram maior número médio de ácaros na folha. Os demais fatores apresentaram o menor nível de ácaros, isso pode ser atribuído ao baixo número de ácaros amostrados no pomar (Figura 8).

Segundo CHABOUSSOU (1987) relata o papel adverso que altas doses de nitrogênio têm sobre o metabolismo celular e sua relação com o aumento da incidência de pragas.

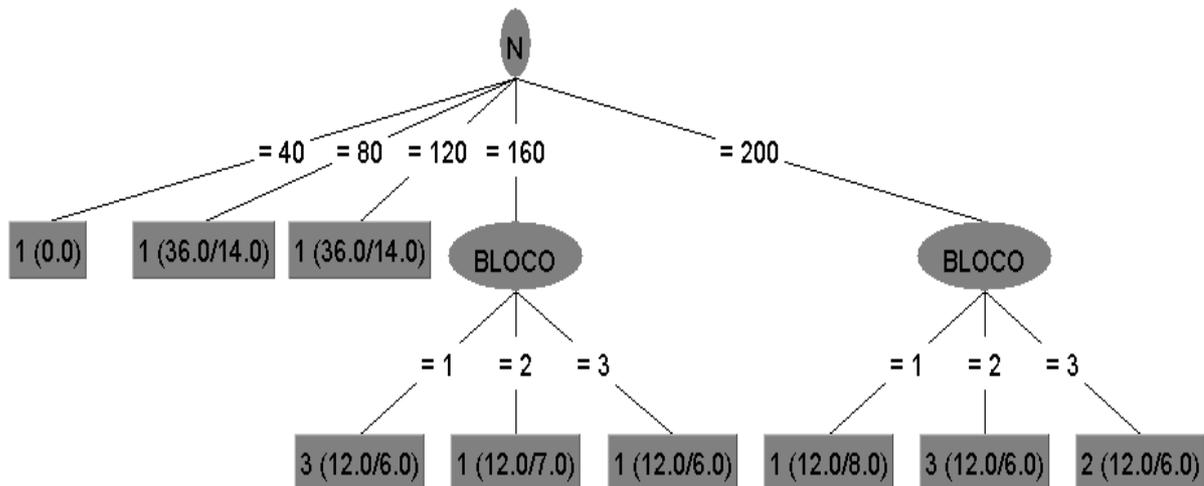


FIGURA 9 - ÁRVORE DE CLASSIFICAÇÃO COM 53,5% DE CONFIABILIDADE ORIGINADA DO TEOR FOLIAR E ADUBAÇÕES DE NITROGENIO E POTASSIO EM RELAÇÃO AO NÍVEL DE ÁCAROS RAJADOS AMOSTRADO NO POMAR DE PÊSSEGO LOCALIZADO EM CAMPO LARGO, PR.

Nos estudos de DAUD, FERES e BUOSI (2007), cita o FLUCKINGER e BRAUN (1999); LARCHER (2000) aonde se afirma que plantas submetidas a condições de estresse (pobreza de solo, poluição atmosférica e micro clima) estão mais predisposta a desenvolver doenças e a apresentar maiores infestações de artrópodes fitófagos.

As análises descritivas e análise de modelos lineares generalizados e Árvore de Classificação estão corroborando os resultados, porém, não esta concordando com os resultados não fundamentados obtidos pela análise de variância.

4 CONCLUSÃO

- a. Pela análise descritiva identifica-se elevada dispersão dos dados;
- b. Os gráficos *Box-plot* são eficientes para identificar pontos discrepantes e apresentar a dispersão das variáveis respostas;
- c. Análise de Variância não é apropriada nas condições deste trabalho para analisar dados de contagem de ácaros.
- d. Nos pomares da Lapa e Campo Largo, os resíduos da ANOVA apresentaram semelhantes dispersões e tendências.
- e. O Modelo Linear Generalizado para Dados Longitudinais, considerando a Distribuição Probabilística de Poisson e Árvore de Classificação, apresentaram-se apropriados para analisar a contagem de ácaros nas folhas de pessegueiro.
- f. O manejo realizado em ambos pomares, não possibilita identificar estrutura de correlação entre as datas de amostragem de ácaros.
- g. Árvore de Classificação é eficiente para classificar 45 amostras de ácaros na folha.
- h. Fator de maior relevância para definir o nível de ácaros rajados no pomar de pessegueiro é a localização, seguido pela concentração de $160 \text{ kg.ha}^{-1} \text{ ano}^{-1}$ de Nitrogênio.
- i. Os modelos estatísticos empregados mostra que os teores foliares e os macros e micronutrientes do solo não influenciaram no número de ácaros.

CONSIDERAÇÕES

Ao concluir este estudo pode-se dizer que a variável oriunda do teor foliar, análise de solo não ajudou explicar a presença ácaros nas folhas. Pode-se ressaltar que ocorreu falha no planejamento do experimento do pomar localizado na Lapa, já que, por meio das análises descritiva é possível verificar que o número de ácaros móveis encontrados no bloco III foi influenciado pelo pomar vizinho, além das unidades experimentais serem acentuadamente heterogêneas e os fatores tratamentos de maior interesse que é o nitrogênio estar localizado nas parcelas principais, diminuindo assim amostra deste.

Por meio dos métodos de análises estatísticas aplicados aos dados de contagem de ácaros, possibilitou concluir que adubação nitrogenada nas menores doses aplicadas influenciaram indiretamente no aumento do número de ácaros. Já para maior dose, observou-se em geral um aumento na variabilidade das respostas.

A técnica de Árvore de Decisão é recomendada para número elevado de observações e uma variabilidade acentuada na variável resposta, quando aplicada nas 30 observações o nível de confiabilidade foi inferior a 50%, porém, com aplicação da técnica de *Bootstrapping* a amostra foi ampliada para 45 observações obtendo um grau de confiança mais que 90%. Portanto a técnica de *bootstrapping* foi relevante para definição do nível de confiabilidade do modelo ajustado por meio da Árvore de Classificação.

REFERÊNCIAS

- AGUIRRE, A. **Uma nota sobre a transformação BOX-COX**. 1997. Disponível em <<http://www.cedeplar.ufmg.br/pesquisas/td/TD%20116.pdf>>. Acesso em Novembro. 2007
- ALVES, V. R. M.; FERREIRA, V. **Algoritmo genético para a solução do problema de roteamento e seqüenciamento de sondas de manutenção**. In: SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL, 38, Goiânia, 2006.
- BANZATTO D. A.; KRONKA, S. do N. **Experimentação agrícola**. Jaboticabal: UNESP, 1989. 53 p.
- BARBIN, D. **Componentes de variância: teoria e aplicações**. Piracicaba: FEALQ, 2004. 120p.
- BENJAMINI, Y. Opening the Box of a Boxplot. **The American Statistician.**, v. 42, n.4 p. 257-262, 1988.
- BOWER, M. KEITH. **Análise de Variância (ANOVA) usando Minitab**. Disponível em <<http://www.minitabbrasil.com.br/novidades/artigos/artigo24.asp>> Acesso em 22 de Dezembro de 2007.
- BUSSAB, W. O.; MORETTIN, P. A. **Estatística Básica**. 5.ed. São Paulo: Editora Saraiva, 2002. 520 p.
- CAMPOS, G. M.; BARBIN, L. E.; SPANÓ, E. C. J.; SILVA, S. R.; PÉCORÁ, D. J. **Estatística Prática para Docentes e Pós-Graduandos**. Capítulo 2. Tipos de variáveis. 2000. Disponível em: http://www.forp.usp.br/restauradora/gmc/gmc_livro/gmc_livro_cap02.html. Acesso em: 10 de Outubro de 2007.
- CARVALHO, D. R.; FREITAS, A. A.; EBECKEN, N. F. F. **A critical review of rule surprisingness measures**. Data Mining IV, Rio de Janeiro: WIT Press, 2003. p.545-556.
- CINCONTI, S., C.; PONTA L.; RABERTO, M.; SCALAS, E. **Poisson-process generalization for the trading waiting-time distribution in a double-auction mechanism**. International Conference on Noise and Fluctuations in Econophysics and Finance.

Proceeding. The International Society for Optical Engineering, Austin, USA, 2005. Disponível em: < <http://cinef.dibe.unige.it/publications/cincotti05b.pdf>>. Acesso em: 10 de Outubro de 2007.

CHABOUSSOU, F. **Plantas doentes pelo uso de agrotóxicos: a teoria da trofobiose**. 2 ed. Porto Alegre: L&PM, 1987. 256 p.

CHOONPRADUB, C.; MCNEIL, D. Can the box plot be improved? **Songklanakarinn Journal of Science and Technology**, Hat Yai, v.27, n.3, p. 649-657, 2005.

COCHRAN, W. G.; COX, G. M. **Experimental designs**. 2 ed. New York: John Wiley & Sons, 1978. 611 p.

COSTA, S. C. **Modelos lineares generalizados mistos para dados longitudinais**. Piracicaba, 2003, (Tese de Doutorado) – ESALQ, Universidade de São Paulo.

DAUD, R. D.; FERES, R. J. F.; BUOSI, R. Ácaros (Arachnida: Acari) Associados a *Bauhinia variegata* L.(Leguminosae) no Noroeste do Estado de São Paulo. **Neotropical Entomology**, Londrina, v. 36, n.2, p. 322 - 325, 2007.

DEMETRIO, C. G. B. **Modelos lineares Generalizados em experimentação agrônômica**. Piracicaba: ESALQ, 2001. 113p.

DRAPER, N. R.; SMITH JR., H. **Applied Regression Analysis**. 3ed. New York: John Wiley & Sons, 1998. 706 p.

EFRON, B; TIBSHIRANI, R. **Cross-Validation and The Bootstrap: Estimating The Error Rate Of A Prediction Rule**. Technical Report (Tr-477), Dept. Of Statistics, Stanford University, 1995.

FADINI, A. M. Spatial distribution of phytophagous mites (Acari: Tetranychidae) on strawberry plants. **Neotropical. Entomology**, Londrina, v. 36, n. 5, p. 783-789, 2007.

FADINI, M; OLIVEIRA, G. H.; VENZON, M.; PALLINI, A.; VILELA, E. F. **Distribuição espacial de ácaros fitófagos (Acari: Tetranychidae) em morangueiro**. *Neotrop. entomol.* , Londrina, v.36, n.5, 2007. Disponível em:<http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1519-566X2007000500021&lng=pt&nrm=iso>. Acesso em: 13 Dezembro 2008.

FACHINELLO, J. C.; COUTINHO, E. F.; MARODIN, G. A.; BOTTON, M.; MAY DE MIO, L. L. **Normas técnicas e documentos de acompanhamento da produção integrada de pêssego**. Pelotas: Signus, 2003. 92p.

FELDENS, M. A. **Descoberta de conhecimento aplicada à detecção de anomalias em base de dados**. Porto Alegre: PPGCC da UFRGS, 1999. 74 p.

FERREIRA, A.; ALVES, L. **Inteligência Artificial e Data Mining**. Salvador, 2005, Dissertação (Mestrado) Universidade Federal da Bahia (UFBA).

FISHER, R.A. **Statistical Methods for Research Workers**. Edinburgh: Oliver e Boyd, 1962.

FLUCKINGER, W.; BRAUN, S. Stress factors of urban trees and their relevance for vigour and predisposition for parasite attacks. **Acta Horticulturae**, Paris, n. 496, p. 325-334. 1999.

GIOLO, S. R. **Análise de Regressão e Correlação**. Curitiba: Universidade Federal do Paraná. 2005.

GOMES, P. F. **Curso de estatística experimental**. 13 ed. Piracicaba: ESALQ-USP, 1990.

HARRISON, T. H. **Intranet Data Warehouse**. São Paulo: Berkeley, 1998. 75p.

HOGG, R.V., LEDOLTER, J. **Applied Statistics for Engineers and Physical Scientists**. New York: Macmillan Publishing Company, 1987. 155 p.

JUNIOR, P. J. R. **Curso sobre o programa computacional R**. Curitiba, Universidade Federal do Paraná (UFPR). 2005. Disponível em: <<http://leg.ufpr.br/Rpira/Rpira/Rpira.html>>. Acesso em: 24 de Dezembro de 2007.

LARCHER, W. **Ecofisiologia vegetal**. São Carlos: Rima, 2000. 531p.

LIANG, K.Y.; ZEGER, S.L. Longitudinal data analysis using generalized linear models. **Biometrika**, Oxford, v.73, n.1, p.13–22. 1986.

MASSAD, E.; SILVEIRA, P. S. P. **Métodos quantitativos: conceitos iniciais**. Disciplina de Métodos Quantitativos em Medicina da Universidade de São Paulo (USP), 2003. Disponível em: <<http://www.dimfm.usp.br/metquant/index.php>> Acesso em: 22 de Dezembro de 2007.

MITCHELL, M. **An introduction to genetic algorithms**. London: A Bradford Book, 1996. 205 p.

MCCULLAGH, P.; NELDER, J. A. **Generalized linear models**. 2 ed. London: Chapman & Hall, 1989. 511 p.

MONTGOMERY, D. C. **Design and analysis of experiments**. 3 ed. New York: John Wiley & Sons, 2001. 856 p.

NOGUEIRA, E. P. F. **Data Mining Usando o SODAS: estudo de Caso Aondenamoro.com**. Rio de Janeiro, 2007, 145 f. Dissertação (Mestrado em Administração e Economia) - Faculdades Ibmecc.

OKU, K.; YANO S.; TAKAFUJI, A. **Different maternal effects on offspring performance in tetranychid mites, Tetranychus kanzawai and T. urticae** (Acari: Tetranychidae) Applied Entomology and Zoology Vol. 37, pp.425-429 (2002)

PÉREZ, L. F.; FREIRE, L. A. C. **Análise de dados longitudinais com aplicação na indústria**. Relatório Técnico, 2003. Universidade Federal do Paraná. <<http://www.est.ufpr.br/rt/fl03.pdf>> Acesso em: 24 de Março de 2007

R DEVELOPMENT CORE TEAM. **R. Writing R extensions** . Disponível em< <http://cran.r-project.org/doc/manuals/R-exts.pdf>>. Acesso em: 23 de Agosto de 2007

R DEVELOPMENT CORE TEAM. **R. A language and environment for statistical Computing**. Disponível em < <http://www.R-project.org>>. Acesso em: 25 de Fevereiro de 2007.

ROHATGI, V. K. **An Introduction to Probability Theory and Mathematical Statistic**. New York: Wiley, 1976 . 684 p.

RYAN, T. A., JOINER, B. L. **Normal Probability Plots and Tests for Normality**. Disponível em: < <http://www.minitab.com/resources/articles/normprob.pdf> >. Acesso em: 26 de Dezembro de 2007.

SIEGEL, S. **Estatística não-paramétrica**. São Paulo: MacGraw-Hill, 1975. 350p.

SILVA, F. C. **Manual de Análises Químicas de Solo, Planta e Fertilizante**. Brasília: Embrapa Solos, 1999. pg 175-222.

SIT, V. **Catalogo f curves for curves fitting**: Handbook 4. Victoria: Ministry of Forest, Forest Science Research Branch, Biometrics information, 1994. XXp.

SNEDECOR, G. W.; COCHRAN, W. G. **Statistical Methods**. 8 ed. Iowa: Iowa State University Press. 1989. Disponível em: <http://www.itl.nist.gov/div898/handbook/eda/section3/eda357.htm>> Acesso em: 26 de Dezembro de 2007.

SPIEGEL, M. R. **Estatística**. 3. ed. São Paulo: Makron Books, 1994. 639 p.

SPSS Inc. **SPSS statistical software**. Version 12.0. Illinois, 2003.

SUSKI, Z. W.; BADOWSKA, T. Effect of the host plant nutrition on the population of the two spotted spider mite, *Tetranychus urticae* Koch (Acarina, Tetranychidae). *Ekologia Polska*, Lomianki, v. 23, n. 1, p. 185-209, 1975.

TOLEDO, G. L.; OVALLE, I. I. **Estatística básica**, 2 ed. São Paulo: Atlas, 1995. 459 p.

VIEIRA, M. R.; CORREA, S. L.; CASTRO, G. M. M. T.; SILVA, S. F. L.; MONTEVERDE, S, M. Effect of papaya (*Carica papaya* L.) cultivated in a protected environment on the occurrence of phytophagous mites and whiteflies. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 26, n. 3, p. 399 – 402, 2004.

WEKALIST. **Interpret M5 model tree output**. Disponível em: <<https://list.scms.waikato.ac.nz/mailman/htdig/wekalist/2003-July/001609.html>>. Acesso em: 21 de Fevereiro de 2007.

WEKA. **MACHINE LEARNING PROJECT**. Disponível em: <<http://www.cs.waikato.ac.nz/~ml>>. Acesso em: 10 de Abril de 2007.

WEKA SOFTWARE. **Weka 3**: Data Mining with Open Source Machine Learning Software in Java. Disponível em: < <http://www.cs.waikato.ac.nz/~ml/weka/index.html> >Acesso em: 10 de abril de 2007.

WITTEN, I. H.; FRANK, E. **Data Mining: Practical machine learning tools with Java implementations**. San Francisco: Morgan Kaufmann, 2000. 369 p.

**CAPÍTULO III - TEORES FOLIARES EM POMARES DE PESSEGUEIRO SOB
ADUBAÇÃO NITROGENADA E POTÁSSICA NA LAPA E
CAMPO LARGO, PR.**

TEORES FOLIARES EM POMARES DE PESSEGUEIRO SOB ADUBAÇÃO NITROGENADA E POTÁSSICA NA LAPA E CAMPO LARGO, PR.

RESUMO

Atualmente estudos realizados em experimentos agrícola, utilizam poucas ferramentas estatísticas, por vezes de forma inadequada, isso pode ser explicado pela diversidade de técnicas estatísticas disponíveis atualmente ou pela complexidade das análises, portanto, deixam-se de se obter importantes resultados. Objetivo é analisar os dados por meio de diferentes métodos estatísticos, verificando a relação entre os teores foliares dos pomares de pêssgo e adubação nitrogenada e potássica. Os experimentos foram conduzidos em pomares de pessegueiro da cultivar Chimarrita na Lapa e Campo Largo, PR, ambos conduzidos no sistema de Produção Integrada de Pêssego (PIP) desde 2003. O experimento foi conduzido em um arranjo de parcelas sub-subdivididas no delineamento em blocos completos casualizados com de três repetições, sendo que no pomar da Lapa foram administradas cinco doses de nitrogênio (40, 80, 120, 160 e 200 kg.ha⁻¹ ano⁻¹ de N), em Campo Largo quatro doses de nitrogênio (80, 120, 160 e 200 kg.ha⁻¹ ano⁻¹ de N), nas sub-parcelas e duas doses de potássio (55 e 110 kg.ha⁻¹ ano⁻¹ de K₂O) nas sub-subparcelas. Foram realizadas análises dos teores foliares (macronutrientes N, P, K, Ca e Mg g.kg⁻¹ e os micronutrientes Fe, Mn, Zn, Cu e B mg.kg⁻¹) em três safras (2003/04, 2004/05 e 2005/06) na Lapa e duas safras (2003/04 e 2004/05) em Campo Largo. Análises estatísticas dos dados realizadas são: Gráfico de dispersão, *Box-plot*, Coeficiente de Variação, Correlação de Spearman, Pearson e Canônica, ANOVA e Modelos Lineares Generalizados. Análises GLM e ANOVA apresentaram as mesmas conclusões. A aplicação de 160 e 200 kg.ha⁻¹ ano⁻¹ de N proporcionou maiores teores de N na folha. A adubação potássica não influenciou nos teores foliares NPK.

Palavra-Chaves Parcelas Sub-Subdividida, Modelos Lineares Generalizados, Pêssego, Nutriente, Nitrogênio, Potássio.

**NUTRITIONAL STATUS OF THE LEAF OF ORCHARDS OF PEACH UNDER
FERTILIZATION NITROGEN AND, POTASSIUM IN LAPA END CAMPO LARGO,
PR.**

ABSTRACT

Currently studies in Agronomy using the minimum possible tools statistics, often inadequately. This can be explained by the variety of statistical techniques currently available or the complexity of the analyses, therefore, leave up to obtain results important. The objective was analyze the data through different statistical methods checking the relationship between the leaf nutritional status of peach orchards and fertilization of nitrogen and potassium. The experiments were conducted in orchards of peach cultivar Chimarrita at Lapa and Campo Largo Country, PR – Brazil, both conducted under the system of Integrated Production of Peach (PIP) since 2003. The experiment was conducted in an arrangement of split-split plots in blocks complete randomized design with three replications, In the orchard of Lapa was applied five nitrogen rate (40, 80, 120, 160 and 200 kg.ha⁻¹ years⁻¹ of N), in Campo Largo four nitrogen rate (80, 120, 160 and 200 kg.ha⁻¹ years⁻¹ of N), in the sub-plots and two rate of potassium (55 and 110 kg.ha⁻¹ years⁻¹ of K₂O) in the sub-subplots. Foliar analyses was performance in three seasons (2003/04, 2004/05 and 2005/06) in Lapa and two seasons (2003/04 and 2004/05) at Campo Largo. Analysis of the data held statistics: Scatterplot, Box-plot, Coefficient of Variation, Correlations Spearman, Pearson and Canonical, Analysis of Variance, Generalized Linear Models. Analysis GLM and ANOVA showed the same conclusions. Fertilization nitrogen 160 and 200 kg.ha⁻¹ year⁻¹, influenced the amount of N found on the sheet. The potassium fertilization did not influence the foliar of levels NPK

Key-words: Split-Split-Plot, Generalized Linear Models, Peach, Nutrient, Nitrogen, Potassium.

1 INTRODUÇÃO

O conhecimento dos teores foliares, com o qual se está trabalhando é de fundamental importância, pois auxiliam na realização de uma avaliação do estado nutricional adequada e racional, a qual, conseqüentemente, irá contribuir para os custos de produção. A quantificação das carências e excessos dos nutrientes presentes é feita em conjunto com análise de solo, avaliação de campo e análise de tecido foliar. Devido à importância dessas análises, os dados contidos nas mesmas devem ser exaustivamente analisados, para tanto, a estatística é uma ferramenta essencial no que tange à análise e interpretação dos resultados.

Contudo, observa-se em trabalhos atuais com teores foliares, que as ferramentas estatísticas são utilizadas minimamente, e por vezes, de forma inadequada. Essa aplicação errônea se deve ao desconhecimento da diversidade de técnicas estatísticas disponíveis atualmente, e também, pela complexidade das análises, diante disso, deixam-se de obter importantes resultados.

O objetivo das ferramentas estatísticas é encontrar, por meio de modelos matemáticos, estruturas que representem o conjunto de dados, no entanto, não há técnica que resolva todos os problemas, pois diferentes métodos têm diferentes objetivos, que apresentam vantagens e desvantagens. Por isso, a importância de se ter o conhecimento do maior número de métodos e seus pressupostos e suas aplicabilidades.

Ao estudar variáveis de forma sistêmica sugere-se procurar a relação causa e efeito entre as variáveis (MATTAR, 2001), desta forma é inevitável pesquisar o grau de relacionamento entre a variável resposta e os fatores que possam influenciar nas conclusões. No entanto, é comum a simplificação das análises estatísticas, que muitas vezes é grosseira, com a finalidade de facilitar a análise de dados ou torná-la viável (NETO, 2006).

Geralmente, as diferenças existentes entre grupos ou populações de interesse, não estão sendo explicadas apenas por uma variável e sim por conjunto delas, podendo esta relação ser explicada por modelos matemáticos (FAYYAD, 1996). O estudo de uma única variável indica que determinado tratamento é mais adequado, porém quando são consideradas outras variáveis relacionadas à variável de interesse, outros tratamentos podem ser mais indicados. Isso pode ser explicado pela interação entre as variáveis e fatores (PONTES, 2005).

Segundo CAMPO, et. al., (2000), é importante a identificação das características e tipo da variável envolvida no experimento. Essas informações irão definir os métodos estatísticos mais adequados e a interpretação correta dos resultados.

O método mais utilizado em análises agronômicas é Análise de Variância (CHRISTENSEN e BEDRICK, 1997), entretanto, ANOVA é facilmente influenciada pela variabilidade acentuada e pela não normalidade dos dados, o que leva, por vezes, a conclusões errôneas. A variabilidade significativa pode ser verificada pelos altos valores dos quadrados médios dos resíduos que estima o afastamento dos dados em relação à média (SIT, 1994). A ANOVA é eficiente para analisar dados contínuos, desde que os pressupostos sejam atendidos, entretanto, para experimentos cujas variáveis respostas são proporções ou contagens, e em geral, não seguem uma distribuição normal, sugere-se, a teoria de modelos lineares generalizados, a qual é acentuadamente difundida (MCCULLAGH e NELDER, 1989; DEMETRIO, 2001).

amostra característica da variável vai determinar qual metodologia a ser usada (ESTELBERGER e REIBNEGGER, 1995; PEARSON, 2001). Outro fator a ser considerado são os métodos de significância sobre os coeficientes obtidos, os quais auxiliam no estudo da existência dessas associações na população da qual a amostra foi extraída e que serviu de base para o cálculo do coeficiente (GIOLO, 2005). Quando o interesse é avaliar o grau de relacionamento entre dois conjuntos de variáveis utiliza-se análise de correlação canônica, procedimento estatístico multivariado que permite o exame da estrutura de relações existentes entre dois grupos de variáveis (TRUGILHO, LIMA e MORI, 2003).

Para evidenciar a importância de se utilizar diversos métodos, o presente trabalho objetivou analisar os dados por meio de diferentes métodos estatísticos, verificando a relação entre os teores foliares e adubação nitrogenada e potássica dos pomares de pêsego da Lapa e Campo Largo.

2 MATERIAL E MÉTODOS

2.1 MATERIAL

Os dados utilizados neste trabalho foram provenientes de dois experimentos em pomares comerciais de pessegueiro (*Prunus persica*) cv. Chimarrita, submetidos a diferentes doses de adubação nitrogenada e potássica, realizados por alunos doutorandos do Programa de Pós-Graduação em Agronomia, área de concentração em Produção Vegetal e componentes do Grupo de Ensino, Extensão e Pesquisa em Produção Integrada – GEEPPI da Universidade Federal do Paraná.

Os ensaios foram conduzidos nas safras de 2003/04 e 2004/05 em Campo Largo, PR e 2003/04, 2004/05 e 2005/06 na Lapa, PR.

O pomar na Lapa localiza-se nas coordenadas: latitude 25°45'35" S, longitude 49°43'47" W e altitude de 908 metros e o clima classificado (Köppen) como Cfb (subtropical úmido), com ocorrência de geada. O pomar é constituído de solo Cambissolo + Neossolo, com afloramentos rochosos e relevo ondulado. Para plantio de pêssego o solo foi escarificado e arado, sendo que a adubação de base com fosfato e homogeneizada com gradagem. A análise do estado nutricional do solo foi realizada por meio de amostragens aleatórias situadas a 0 – 20 cm de profundidade, 20 cm do tronco (PT) e um terço da projeção da copa (PC), em cada parcela. As amostras, para análise química de rotina, foram coletadas em 14/07/2004 e encaminhadas a Fundação ABC, Castro (Pr) (Tabela 1). A variedade Chimarrita foi plantada no espaçamento 5,0 x 2,5m e a condução foi em taça.

TABELA 1 - ANÁLISE QUÍMICA DO SOLO EM POMAR DE PESSEGUEIRO NO PERIMETRO DO TRONCO (PT), NA PROJEÇÃO DA COPA (PC), A 0 - 20 CM DE PROFUNDIDADE, EM 2004 (LAPA - PR).

Local	Prof. ¹	P resina	M.O.	pH	H+Al ⁺³	Al ⁺³	K ⁺	Ca ⁺²	Mg	SB	CTC	V
		mg/dm ³	g/dm ³	CaCl ₂	-----mmolc/dm ³ -----							
PC	00 - 20	19,9	31,5	6,6	15,3	0,0	3,4	97,9	55,5	156,7	172,0	90,5
PT	00 - 20	38,3	31,6	6,6	16,8	0,0	3,7	99,2	43,5	146,4	163,2	89,4
Média		29,1	31,6	6,6	16,1	0,00	3,5	98,5	49,5	151,6	167,6	89,9

NOTAS: ¹Prof. Profundidade da coleta em cm. M.O. Matéria orgânica

O pomar na Lapa localiza-se nas coordenadas: latitude 25°45'35" S, longitude 49°43'47" W e altitude de 908 metros e o clima classificado (Köppen) como Cfb (subtropical úmido), com ocorrência de geada. Na figura 1 é apresentado as temperatura (°C) máximas, mínimas e média entre os anos de 2003 a 2006.

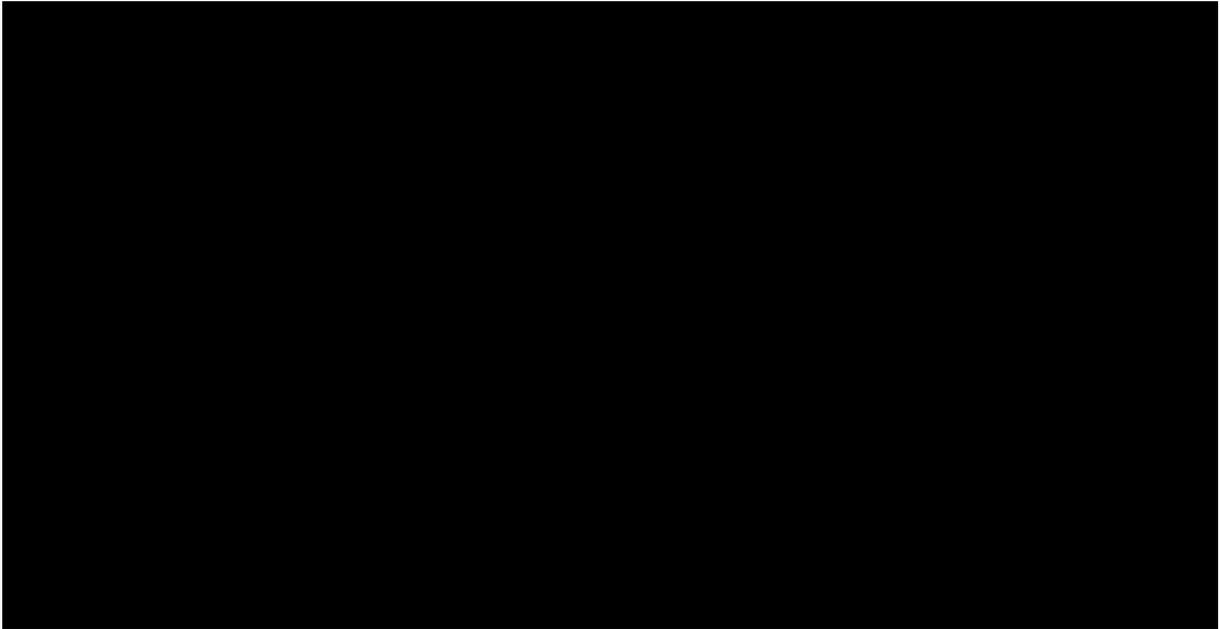


FIGURA 1 - DADOS DE TEMPERATURA MÁXIMA, MÍNIMA E MÉDIA DIÁRIA (°C), DA ESTAÇÃO METEOROLÓGICA DO SIMEPAR MAIS PRÓXIMA DAS ÁREAS EXPERIMENTAIS

FONTE: SIMEPAR - Sistema Meteorológico do Paraná

O pomar é constituído de solo Cambissolo; o pomar foi escarificado e arado durante a instalação, simultaneamente a adubação de base foi fosfato. A análise de solo, nas profundidades de 0 - 20 cm e 20 - 40 cm foi, realizada na implantação do experimento. As amostras, para análise química de rotina, foram encaminhadas a Fundação ABC, Castro (Pr) (Tabela 2). A variedade Chimarrita foi plantada no espaçamento 4,0 x 1,25m e condução foi em Y.

TABELA 2 - ANÁLISE QUÍMICA DO SOLO EM POMAR DE PESSEGUEIRO NO PERÍMETRO DO TRONCO (PT), NA PROJEÇÃO DA COPA (PC) E NA ENTRE LINHA (EL), A 0 - 20 E 20 - 40 CM DE PROFUNDIDADE, EM 2003 (CAMPO LARGO - PR).

Local	Prof. ¹	pH	Al ⁺³	H+Al ⁺³	Ca ⁺² +Mg ⁺²	Ca ⁺²	K ⁺	T	P	C	pH	V
		CaCl ₂	-----cmol/dm ³ -----							mg/dm ³	g/dm ³	SMP
PT	00 - 20	4,9	0,2	5,8	4,39	3,4	0,4	10,6	21,3	27,5	5,80	45,44
PT	20 - 40	4,7	0,5	6,7	3,56	2,5	0,1	10,4	7,9	20,8	5,60	35,33
PC	00 - 20	5,3	0,0	4,6	6,33	4,6	0,4	11,3	9,1	26,9	6,10	59,29
PC	20 - 40	5,0	0,0	5,4	4,10	2,7	0,1	9,63	3,7	23,2	5,90	43,93
EL	00 - 20	5,1	0,0	5,0	5,05	3,6	0,3	10,3	6,1	27,5	6,00	51,60
EL	20 - 40	5,1	0,0	5,0	4,16	2,6	0,1	9,3	4,2	19,6	6,00	46,24
Média		5,0	0,1	5,4	4,60	3,2	0,2	10,3	8,72	24,25	5,90	46,97

NOTAS: ¹Prof. Profundidade da coleta em cm.

O delineamento experimental, devido às condições de relevo dos pomares, foi em três blocos ao acaso, em ambos os pomares, onde os tratamentos constituíram-se de diferentes doses de nitrogênio (N) e potássio (K). Na Lapa, os tratamentos foram compostos por cinco doses de nitrogênio (40, 80, 120, 160 e 200 kg.ha⁻¹ ano⁻¹) e duas de potássio (55 e 110 kg.ha⁻¹ ano⁻¹). Em Campo Largo, o experimento foi conduzido com quatro doses de nitrogênio (80, 120, 160 e 200 kg.ha⁻¹ ano⁻¹) e duas de potássio (55 e 110 kg.ha⁻¹ ano⁻¹). O croqui das áreas experimentais é apresentado na figura 1.

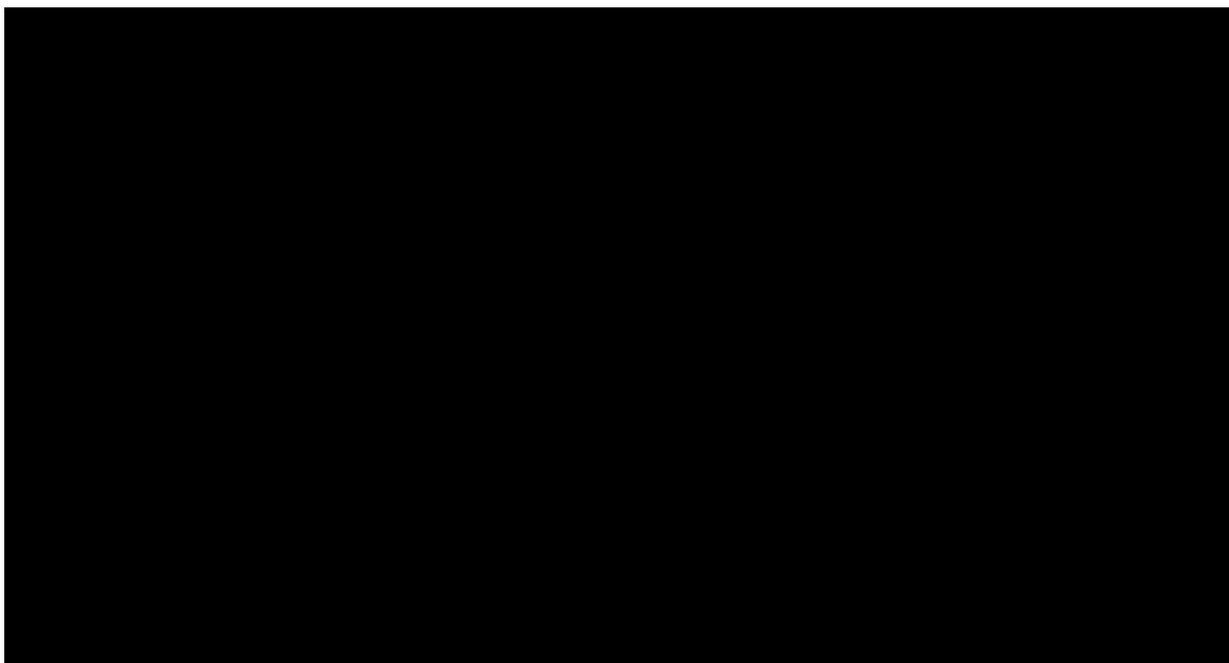


FIGURA 2 - DISPOSIÇÃO DE PÊSSEGO EM SUBPARCELAS E SUB-SUBPARCELAS CONSTITUÍDA DE DOSES DE NITROGÊNIO E POTÁSSIO EM DOIS POMARES, SITUADOS NA LAPA E CAMPO LARGO, PR, 2003 a 2006.

O nitrogênio foi aplicado manualmente na forma de uréia, com um dosador previamente calibrado, na projeção da copa (FACHINELLO, et. al., 2003), sendo que 30% foram aplicados no início da brotação, 30% no raleio e 40% após a colheita. O potássio foi aplicado na forma de cloreto de potássio, sendo 60% aplicado no início da brotação e o restante após o raleio. A aplicação foi feita na projeção da copa, juntamente com o nitrogênio.

A avaliação nutricional dos pessegueiros foi realizada por meio dos teores foliares macronutrientes (N, P, K, Ca e Mg g.kg⁻¹) e os micronutrientes (Fe, Mn, Zn, Cu e B mg kg⁻¹). Foram coletadas, na 14ª semana após plena floração, 80 folhas completas (limbo + pecíolo) por sub-sub-parcelas, localizadas na parte mediana dos ramos do ano, a 1,7 m de altura e em todo o redor da planta. As coletas foram feitas nas safras de 2003/04 e 2004/05 em

Campo Largo e 2003/04, 2004/05 e 2005/06 na Lapa. As folhas foram secas em estufa com circulação forçada de ar a $60 \pm 2^\circ\text{C}$ até massa constante, e as amostras foram encaminhadas para o Laboratório de Nutrição Mineral de Plantas, da Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA Clima temperado), Pelotas, RS e analisada segundo metodologia descrita por SILVA (1999).

2.2 MÉTODOS

Os dados dos teores foliares foram submetidos análises práticas e descritivas.

Entretanto, para este estudo aprofundado, devido à quantidade de análises geradas pelos macro e micronutrientes (N, P, K, Ca, Mg g.kg^{-1} e Fe, Mn, Zn, Cu, B mg kg^{-1} respectivamente) foram escolhidas as variáveis que mais responderam aos tratamentos de adubação nitrogenada e potássica. Portanto, somente as variáveis Nitrogênio, Fósforo e Potássio (NPK) foram submetidas às análises de correlações de Spearman, Pearson, Análise de Variância e Análise de Modelos Lineares.

Os pressupostos da foram verificados por meio do Teste de Shapiro Wilk e Teste Bartlett e pelos gráficos obtidos por *default* do *software R*. (*R Development Core Team* 2005). Para comparações de médias foi aplicado o teste Tukey com 5% ($p\text{-valor} < 0,05$) de significância (MONTGOMERY, 2001; BARBIN, 2004).

Para explicar a influência dos tratamentos sobre o nível de ácaros, considerando a função da Distribuição Probabilística de Gauss utilizou-se Modelos Lineares Generalizados (MCCULLAGH e NELDER, 1989; DEMETRIO, 2001).

Para verificar a dependência entre as doses dentre os elementos e temperatura foi utilizada a Análise de Correlação de Pearson e para relacionar as doses de adubação nitrogenada e potássica com os teores foliares de macro e micronutrientes utilizou-se a Análise de Correlação de Spearman (SIEGEL, 1975; SPIEGEL, 1994; PEARSON, 2001; GIOLO, 2005).

As análises estatísticas aplicadas aos dados foram realizadas pelos *softwares R* (*R Development Core Team* 2005), SPSS e STATGRAPHICS Centurion.

3 RESULTADOS E DISCUSSÃO

Linearizando a apresentação dos resultados, segue-se uma seqüência de análises de modo que permita facilmente a comparação dos resultados obtidos pelos diferentes métodos aplicados a uma mesma variável.

Para ambos os pomares, inicialmente, optou-se por uma metodologia estatística descritiva, onde foram analisadas graficamente as dispersões dos teores foliares dos macro e micronutrientes, abordando as interpretações teóricas. Em seguida são apresentados os coeficientes de variação, os quais descrevem a dispersão dos elementos dentre os tratamentos, finalizando com a análise do *Box-plot*, que aborda individualmente cada safra e posteriormente em conjunto.

Após a análise descritiva, verificou-se a associação entre os teores foliares por meio do Teste de Correlação de Pearson e a relação entre os tratamentos e safras com teores foliares por meio do Teste de Correlação de Spearman e Coeficiente de Contingência “C”. A relação entre as safras foi verificada por meio de Análise de Correlação Canônica. Após a verificação dos pressupostos do modelo, apresentou-se a Análise de Variância e Comparação de Médias e em seguida, a Análises de Modelos Lineares Generalizados.

3.1 ANÁLISE DE DISPERSÃO

Lapa. Os valores foram comparados com teor estabelecidos pela Comissão de Fertilidade do Solo do RS/SC. Portanto, foram encontrados, em níveis abaixo do normal, Mg na safra 2003/04, Ca nas safras de 2003/04 e 2004/05 e de N nas safras 2004/05 e 2005/06. Em todos os anos o P se encontrou em níveis normais. Níveis normais de Ca foram encontrados na safra de 2005/06 e de Mg nas safras de 2004/05 e 2005/06. O K apresentou-se acima do normal em todas as safras. Com exceção das variáveis K na safra de 2003/04 e N nas safras de 2004/05 e 2005/06, os demais nutrientes, não apresentaram variabilidade acentuada dentre os tratamentos de adubação nitrogenada. Não foi possível identificar dados discrepantes que possam influenciar nas análises estatísticas. (FIGURA 2).

Em geral, o fator safra também teve pequena variação no teor dos macronutrientes.

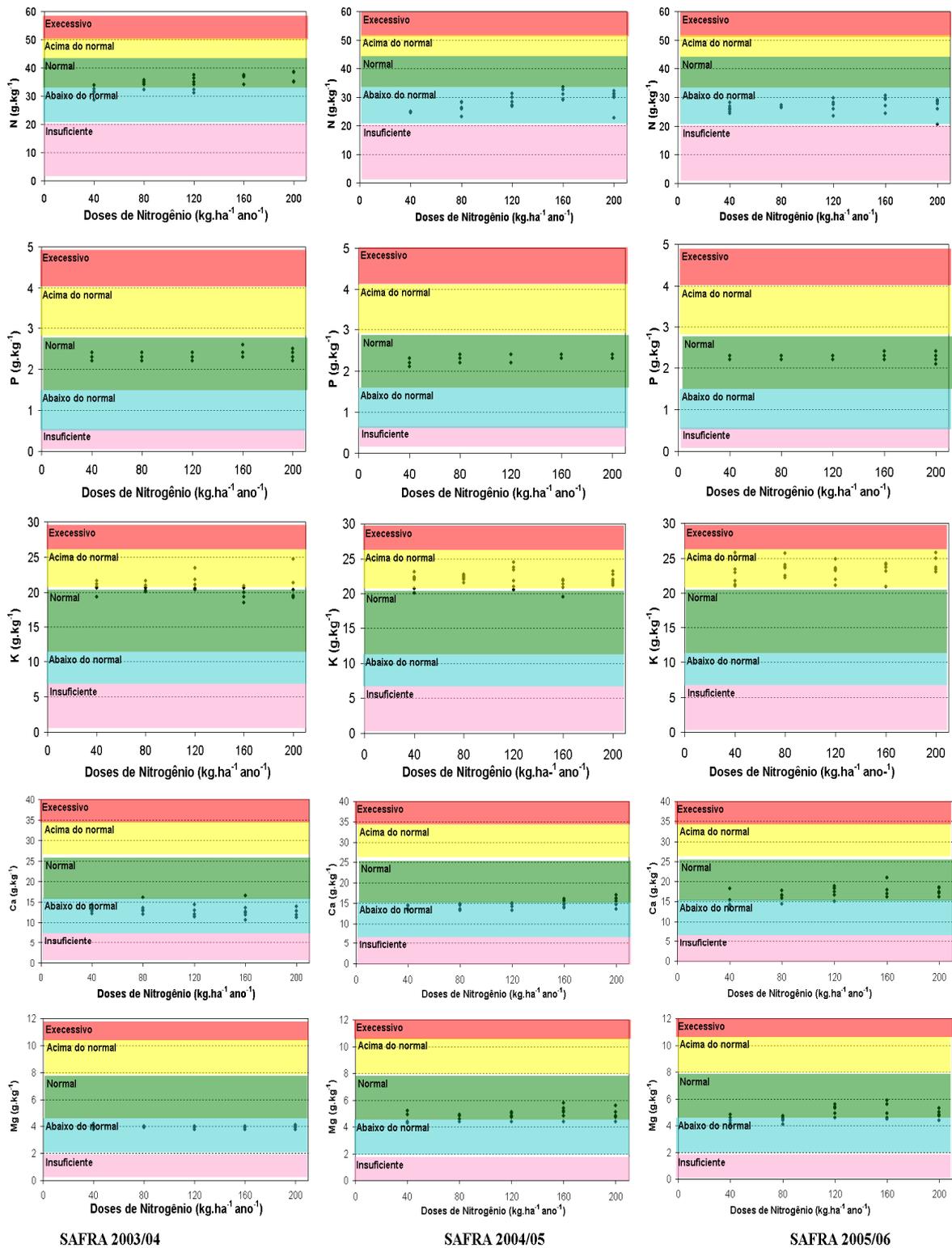


FIGURA 3 - DISPERSÃO DE MACRONUTRIENTES, OBSERVADOS EM DIAGONE FOLIAR, NA PRESENÇA DE ADUBAÇÃO NITROGENADA ($\text{Kg.ha}^{-1} \text{ano}^{-1}$), APLICADA EM POMARES DE PÊSSEGO 'CHIMARRITA'. LAPA, PR. NAS SAFRAS 2003/04, 2004/05 E 2005/06.

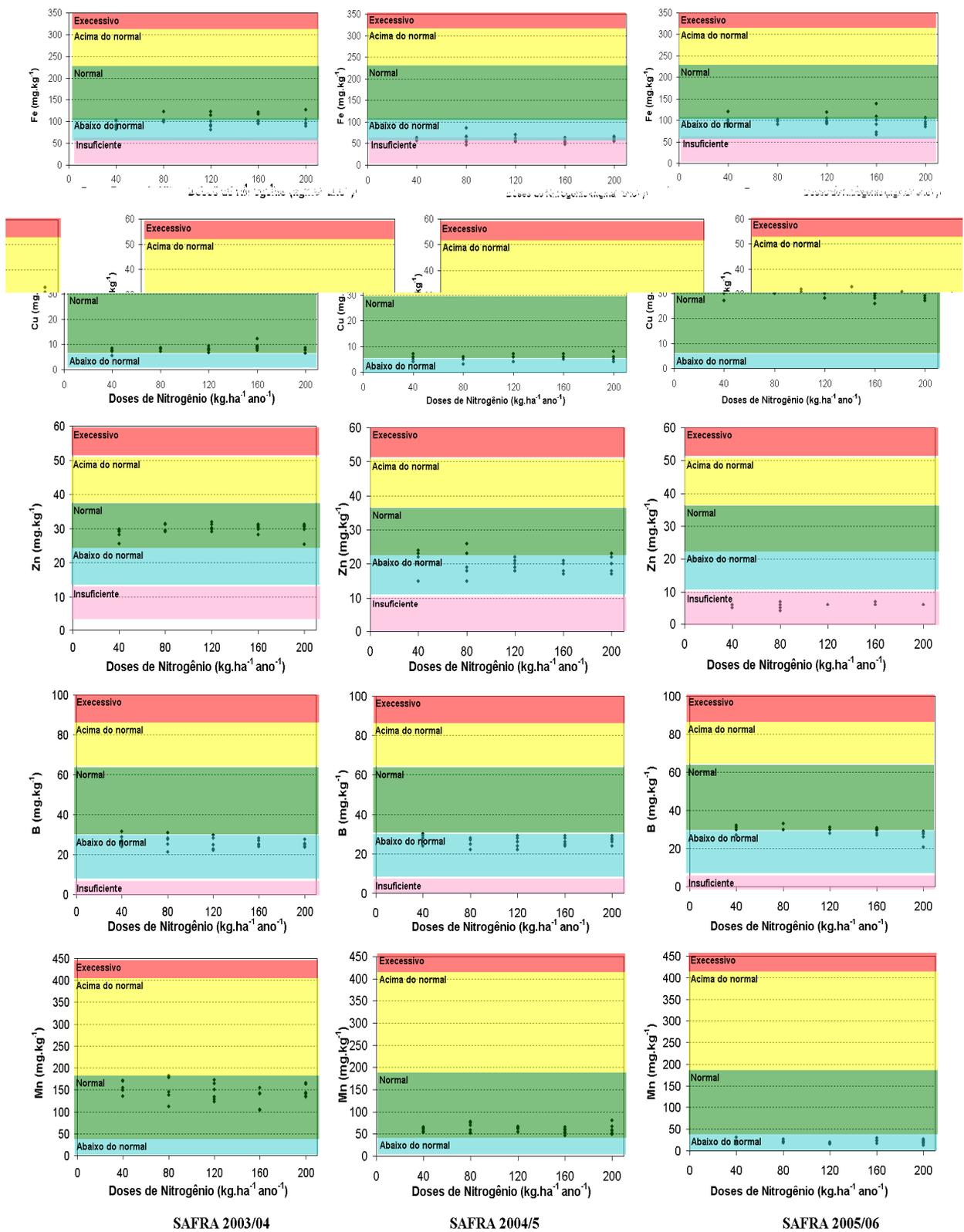


FIGURA 4 - DISPERSÃO DE MICRONUTRIENTES, OBSERVADOS NOS TEORES FOLIARES, NA PRESENÇA DE ADUBAÇÃO NITROGENADA ($\text{kg.ha}^{-1} \text{ano}^{-1}$), APLICADA EM POMARES DE PÊSSEGO 'CHIMARRITA'. LAPA, PR. SAFRAS 2003/04, 2004/05 E 2005/06.

Os dados revelaram que houve redução dos elementos Mn e Zn entre as safras de 2003/04 e 2005/06. Com exceção das variáveis Fe e Zn, nas safras de 2004/05 e 2005/06 e Mn na primeira safra, as demais variáveis não apresentaram, visualmente, elevada variabilidade dentre as doses de adubação nitrogenada (FIGURA 3).

Os macros e micros nutrientes na presença das doses de nitrogênio não apresentaram dados com dispersões acentuadas, portanto, indicando que provavelmente o valor do quadrado médio dos resíduos não seja elevado, ou seja, a distância dos dados em relação à média não será alta, mantendo-se na mesma faixa de interpretação (FIGURA 2 e 3).

Campo Largo. Em ambas as safras N, K e Ca se encontraram em níveis normais. Somente Mg se encontrou em níveis abaixo do normal na safra de 2003/04. A maior dispersão foi verificada para o K em ambas as safras, porém de maneira geral, nas safras de 2003/04 e 2004/05, não houve dispersão acentuada, portanto, não há indício de heterocedasticidade (Figura 4). Ainda, não se observa o efeito da adubação nitrogenada sobre os níveis de macronutrientes.

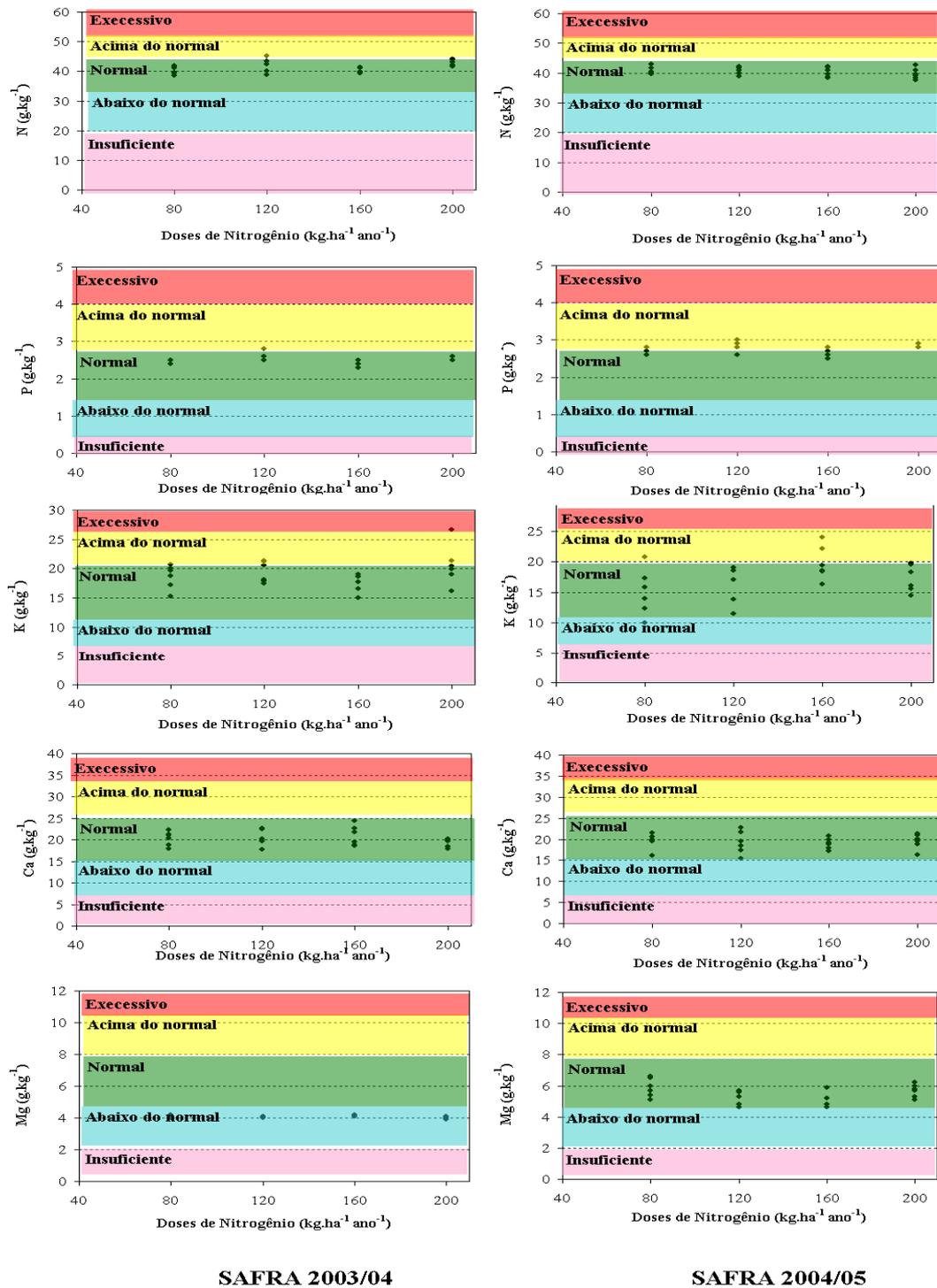


FIGURA 5 - DISPERSÃO DE MACRONUTRIENTES, OBSERVADOS NOS TEORES FOLIARES, NA PRESENÇA DE ADUBAÇÃO NITROGENADA ($\text{kg}\cdot\text{ha}^{-1}\text{ ano}^{-1}$), APLICADA EM POMARES DE PÊSSEGO 'CHIMARRITA'. CAMPO LARGO, PR. SAFRAS 2003/04 E 2004/05.

Na Figura 5 pode ser observado abaixo do normal, o Zn e o Mn, estavam acima do normal na safra de 2004/05, e os demais micronutrientes se encontraram em níveis normais em ambas às safras. Os elementos Zn e Mn na segunda safra foram o que apresentaram maior dispersão dos dados.

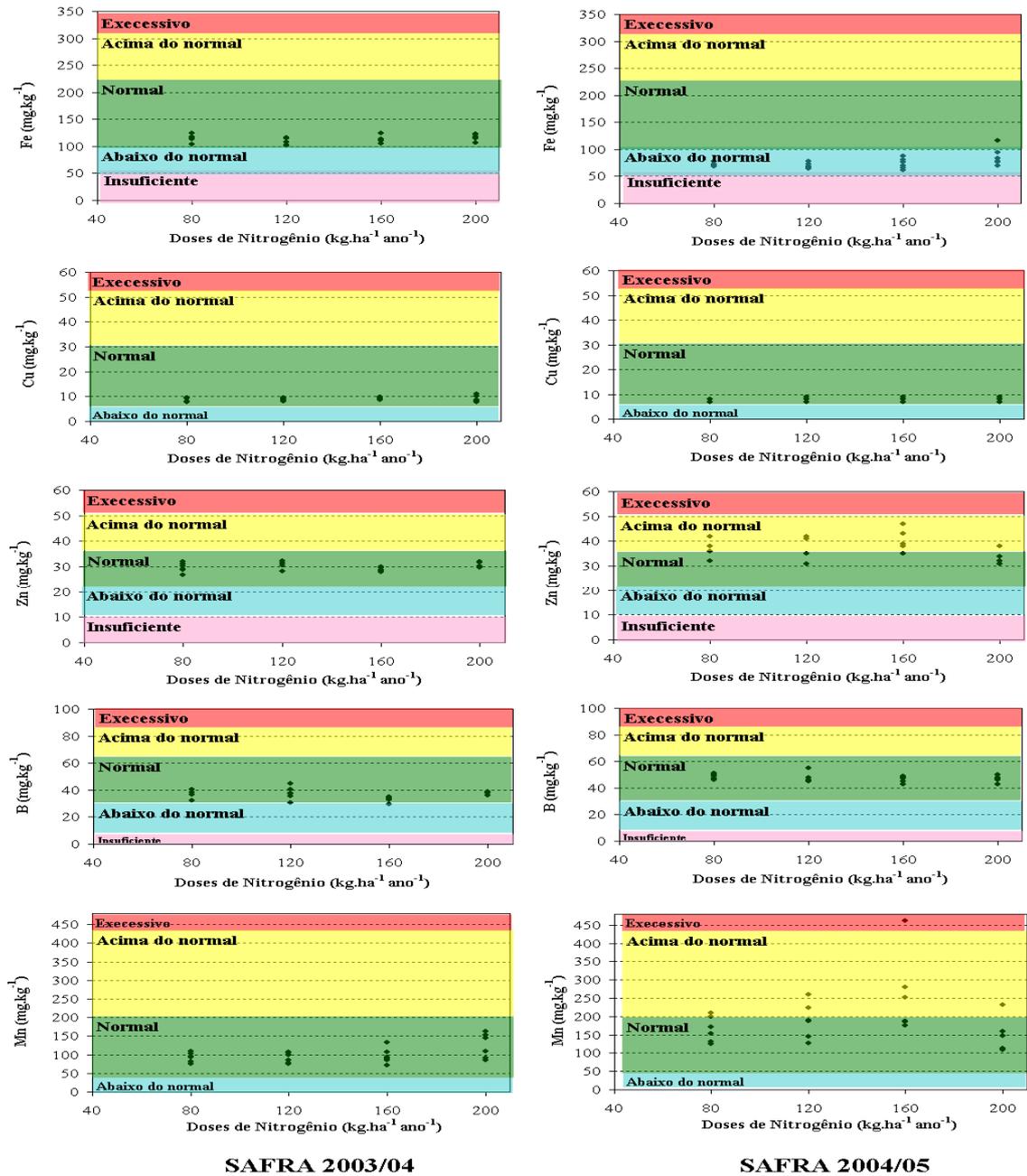


FIGURA 6 - DISPERSÃO DE MICRONUTRIENTES, OBSERVADOS NOS TEORES FOLIARES, NA PRESENÇA DE ADUBAÇÃO NITROGENADA (kg .ha⁻¹ ano⁻¹), APLICADA EM POMARES DE PÊSSEGO 'CHIMARRITA'. CAMPO LARGO, PR. SAFRAS 2003/04 E 2004/05.

Sob ponto de vista nutricional os pomares de Campo Largo e Lapa se apresentavam com plantas adequadamente nutridas (Comissão de Fertilidade, 2004) (Figura 1, 2, 3 e 4). Resultados semelhantes foram encontrados por MATTOS, FREIRE e MAGNANI (1991) em pomares de pêssigo cv. Diamante, onde foram ministradas 5 doses de nitrogênio, dentre 135 a 675 g/planta; e por DOLINSK, et. al. (2005) em pomares de pêssigo cv. Chimarrita, com três tratamentos de nitrogênio (40, 80 e 160 kg.ha⁻¹ ano⁻¹). Nos estudos realizados por SCHUHMACHER, (1997) em que avalia teores foliares de milho (*Zea mays L.*), afirma que os níveis sugeridos por diversos autores se reportam a aos teores foliares ideais, portanto valores com baixa oscilação em relação aos valores sugeridos pelos autores, não caracterizam necessariamente deficiência nutricional.

Além da verificação gráfica da dispersão dos dados, também é possível, por meio do coeficiente de variação (CV), constatar a grandeza da dispersão, permitindo comparar os resultados deste trabalho com estudos com as mesmas variáveis de interesse.

As estatísticas de dispersão dos dados oriundos dos teores foliares de macro e micronutrientes estão apresentadas na Tabela 1 e 2, onde podem ser observados os CV's.

3.2 COEFICIENTES DE VARIAÇÃO

Macronutrientes da Lapa e Campo Largo. No geral, os coeficientes de variação obtidos foram baixos, confirmando a ausência de *outlier* e dispersão não acentuada na figura 1. No pomar da Lapa, a menor dispersão foi verificada para o P em todas as safras, o que corrobora com os resultados obtidos e apresentados na Figura 1. As variáveis dos macronutrientes, N nas safras 2004/05 e 2005/06, Ca nas safras de 2003/04 e 2004/06 e Mg na safra 2005/06 apresentaram maior dispersão (Tabela 1), resultado que não é possível verificar no gráfico de dispersão (Figura 1). No pomar de Campo Largo, o K foi o que apresentou maior dispersão, sendo que os demais elementos não apresentaram dispersão acentuada (Tabela 1).

TABELA 3 – COEFICIENTE DE VARIAÇÃO DOS MACRONUTRIENTES, NOS DIFERENTES DOSES DE ADUBAÇÕES NITROGENADAS REALIZADAS NAS SAFRAS DE 2003/04, 2004/05 e 2005/06 EM POMARES DE PÊSSEGO ‘CHIMARRITA’. LAPA E CAMPO LARGO, PR.

Dose de N kg.ha ⁻¹	SAFRAS														
	*3/4	4/5	5/6	3/4	4/5	5/6	3/4	4/5	5/6	3/4	4/5	5/6	3/4	4/5	5/6
	N (g.kg ⁻¹)			P (g.kg ⁻¹)			K (g.kg ⁻¹)			Ca (g.kg ⁻¹)			Mg (g.kg ⁻¹)		
-----%-----															
Pomar da Lapa															
40	5	1	5	3	3	2	4	5	8	6	4	12	4	8	7
80	4	7	2	3	4	2	3	2	5	10	5	7	1	4	5
120	7	6	8	3	4	2	6	7	6	9	5	8	2	6	8
160	4	6	8	5	2	4	4	4	5	16	6	12	2	9	11
200	5	12	12	5	2	5	10	3	4	8	8	5	3	8	6
Total	7	10	8	4	4	3	6	5	6	10	7	11	2	8	10
Pomar de Campo Largo															
80	3	3		2	3		11	25		8	10		2	10	
120	6	3		5	5		9	19		10	14		1	9	
160	2	4		3	5		9	14		11	7		1	9	
200	2	5		2	2		17	13		5	9		2	7	
Total	4	4		4	5		13	20		9	10		2	10	

* 3/4 = Ano de 2003 para 2004; 4/5 = Ano de 2004 para 2005; 5/6 = Ano de 2005 para 2006

Micronutrientes da Lapa e Campo Largo. A maior dispersão observada, no pomar da Lapa, foi para o Cobre na safra 2004/05 e para o Mn em 2005/06 (Tabela 2). No pomar de Campo Largo, com exceção do Boro, na safra de 2003/04 e do Ferro na safra de 2004/05, as demais variáveis não apresentaram variabilidade acentuada dentre os tratamentos de adubação nitrogenada, corroborando com os coeficientes de variação (Tabela 2). As variações no teor de Mn são comuns, dado a sensibilidade deste elemento as variações do pH e potencial de oxiredução (MOTTA et. al, 2007).

Importante considerar que foram encontrados resultados semelhantes aos pomares de pessegueiro da Lapa e Campo Largo nas mesmas condições, avaliado por DOLINSKI et. al. (2005), o qual mostra que os micronutrientes apresentaram maior dispersão que os macronutrientes e que não houve influência das doses de nitrogênio (40, 80 e 160 kg.ha⁻¹ ano⁻¹), aplicados no solo em forma de uréia.

No estudo sobre a influência da aplicação do calcário nos teores foliares de S, Cu, Fe, Mn e Zn no solo e na planta de milho (ASSMANN, 1995) encontrou valores que foram, na maioria, superiores aos encontrados nos pomares da Lapa e Campo Largo.

TABELA 4 - COEFICIENTES DE VARIAÇÃO (CV) DOS MICRONUTRIENTES, NOS DIFERENTES DOSES DE ADUBAÇÕES NITROGENADAS REALIZADAS NAS SAFRAS 2003/04, 2004/05 e 2005/06 EM POMAR DE PÊSSEGO 'CHIMARRITA' LAPA E CAMPO LARGO, PR.

Dose de N kg.ha ⁻¹	Safras														
	*3/4	4/5	5/6	3/4	4/5	5/6	3/4	4/5	5/6	3/4	4/5	5/6	3/4	4/5	5/6
	Fe (mg.kg ⁻¹)			Cu (mg.kg ⁻¹)			Zn (mg.kg ⁻¹)			B (mg.kg ⁻¹)			Mn (mg.kg ⁻¹)		
-----%-----															
Pomar da Lapa															
40	9	4	13	13	20	6	5	15	7	10	9	6	10	8	26
80	9	21	4	6	23	4	4	20	18	12	9	4	18	18	15
120	16	10	9	11	18	5	3	8	0	13	10	5	14	8	9
160	11	9	27	17	13	5	4	9	8	6	7	5	17	12	21
200	13	7	9	10	25	8	7	13	0	6	7	12	9	19	24
Total	12	12	14	14	19	6	5	13	10	10	8	8	14	14	20
Pomar de Campo Largo															
80	6	5		9	5		6	9		8	4		13	21	
120	6	7		6	8		4	14		13	7		15	26	
160	6	13		5	9		3	11		6	5		22	42	
200	5	19		14	11		3	9		3	5		27	30	
Total	6	15		9	8		5	12		10	5		24	39	

* 3/4 = Ano de 2003 para 2004; 4/5 = Ano de 2004 para 2005; 5/6 = Ano de 2005 para 2006

As análises de realizadas verificando os CV's ajudam a confirmar as análises realizadas com gráficos de dispersão. A vantagem das análises de dispersão sobre os CV's é que possibilita instantaneamente a interpretação prática, entretanto é mais robusta. Por outro lado os resultados do CV's são exatos, possibilitando comparar o comportamento variável a outros experimentos semelhantes.

Outro, porém, baseado nestas análises, quando a variabilidade dos dados não é elevada, não é possível predizer se os pressupostos da ANOVA serão atendidos.

Após apresentação da dispersão dos dados, também se fez necessário identificar a existência de pontos "outliers". As observações que apresentam uma distância acentuada dos demais pontos observados são conhecidas como "outliers" ou dados discrepantes. Estes pontos podem influenciar significativamente as análises dos dados, conseqüentemente, obtendo conclusões que não representam os dados mensurados (SCHWERTMAN E SILVA, 2006). TUKEY (1997) sugere um método simples que identifica graficamente os pontos discrepantes, o *Box-plot*, que identifica os pontos discrepantes, fornecendo o resumo dos cinco números (mínimo ($Q_1 - 1,5(Q_3 - Q_1)$), 1º quartil (Q_1), mediana ou 2º quartil (Q_2), 3º quartil (Q_3) e o valor máximo ($Q_1 + 1,5(Q_3 - Q_1)$) dos dados), permitindo comparar as variáveis

da mesma natureza em uma mesma escala dentre as doses de nitrogênio (BUSSAB e MORETTIN, 1987).

Portanto a seguir é apresentada análise para identificar pontos “*outliers*” que possam influenciar sensivelmente nas medidas de tendência central.

3.3 ANÁLISE DE DADOS DISCREPANTES

Lapa. Os pontos discrepantes, para os macronutrientes, podem ser observados nas seguintes variáveis: N com três observações na aplicação de $40 \text{ kg ha}^{-1} \text{ ano}^{-1}$ na safra de 2003/04 e na aplicação de $200 \text{ kg ha}^{-1} \text{ ano}^{-1}$ nas safras de 2004/05 e 2005/06; P com um dado em $160 \text{ kg ha}^{-1} \text{ ano}^{-1}$ de N na safra de 2003/04, K com três observações, na safra de 2003/04 nas doses de 40 e $200 \text{ kg ha}^{-1} \text{ ano}^{-1}$ de N, Ca com quatro dados na safra de 2003/04 em 80 e $160 \text{ kg ha}^{-1} \text{ ano}^{-1}$ de N e na safra de 2005/06 de $40 \text{ kg ha}^{-1} \text{ ano}^{-1}$ de N, Mg com uma observação no tratamento de $40 \text{ kg ha}^{-1} \text{ ano}^{-1}$ de N na safra de 2003/04. Os pontos discrepantes foram observados em uma população de 90 observações (Figura 6).

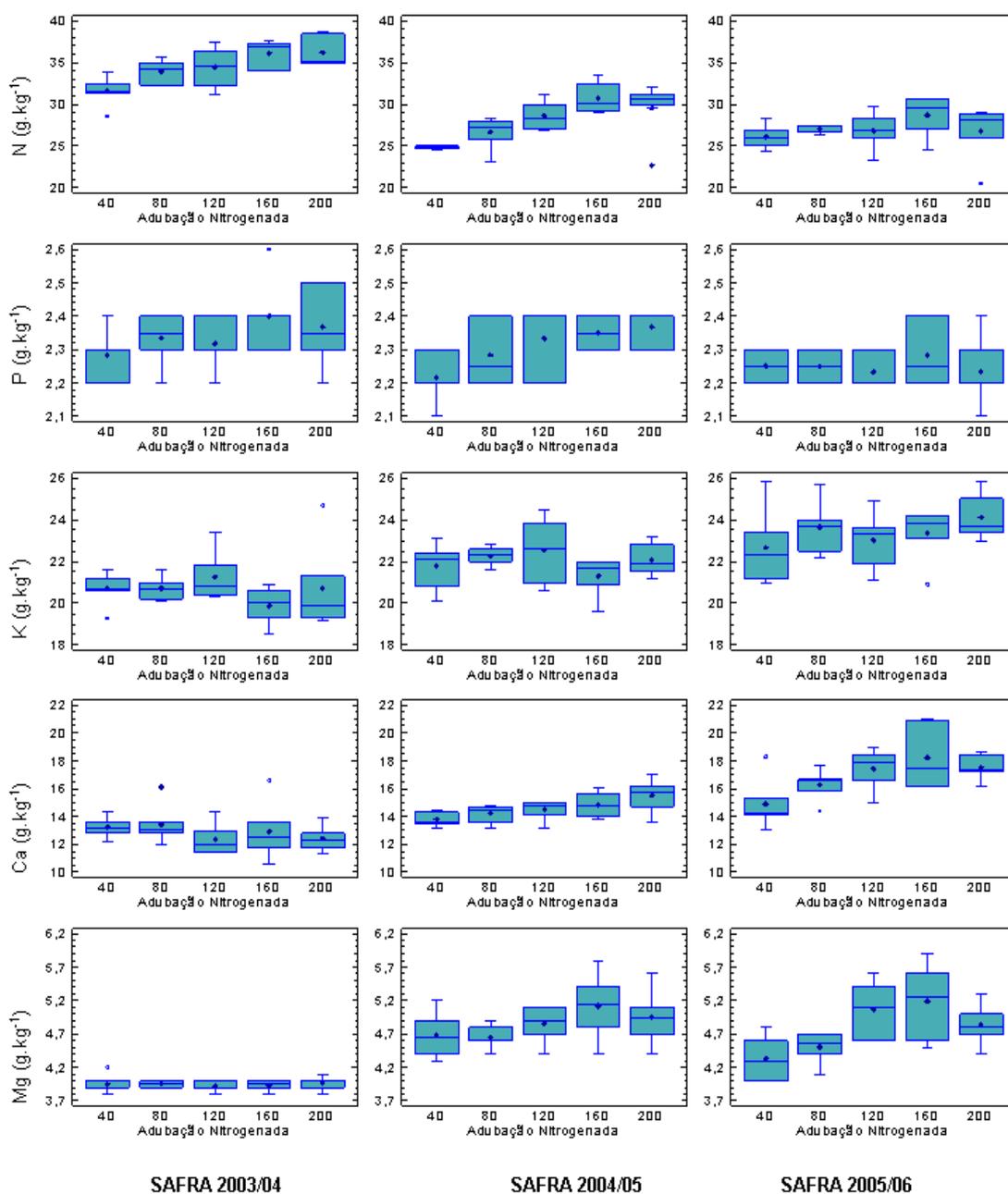


FIGURA 7 - *BOX-PLOT* DOS MACRONUTRIENTES, NOS DIFERENTES DOSES DE ADUBAÇÕES NITROGENADAS ($\text{kg ha}^{-1} \text{ano}^{-1}$) REALIZADAS NO POMAR DE PESSEGUEIROS 'CHIMARRITA'. LAPA, PR.

No mesmo universo de 90 dados, dentre os micronutrientes, o Fe apresentou quatro dados discrepantes, (na safra de 2003/04 nas doses 80 e 200 $\text{kg ha}^{-1} \text{ano}^{-1}$ de N; em 2004/05, na concentração de 80 $\text{kg ha}^{-1} \text{ano}^{-1}$ de N; em 2005/06 para 40 $\text{kg ha}^{-1} \text{ano}^{-1}$ de N) (Figura 7).

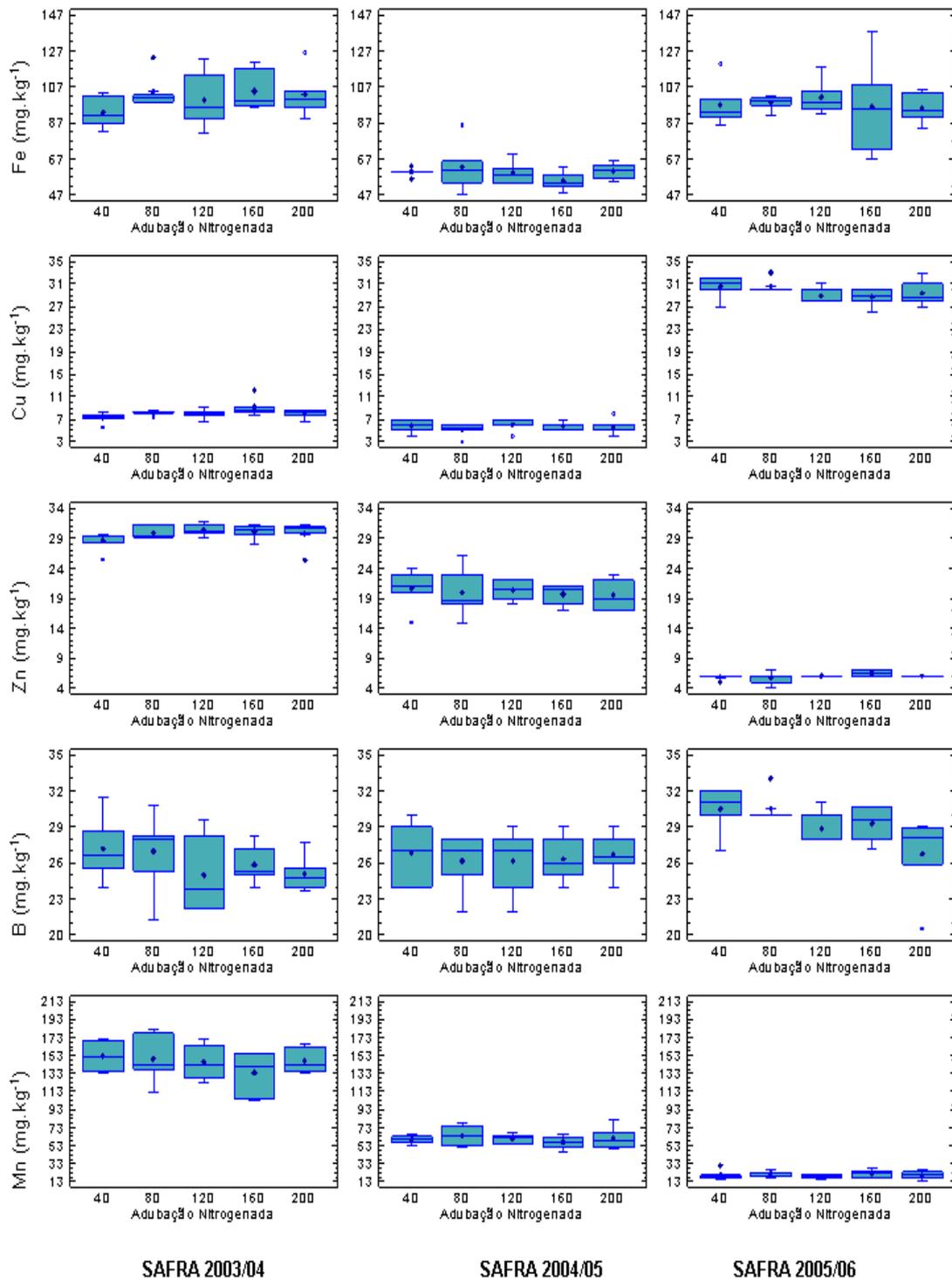


FIGURA 8 - *BOX-PLOT* DOS MICRONUTRIENTES, NOS DIFERENTES DOSES DE ADUBAÇÕES NITROGENADAS ($\text{kg ha}^{-1} \text{ ano}^{-1}$) REALIZADAS EM POMAR DE PÊSSEGO 'CHIMARRITA'. LAPA, PR.

O elemento Cu observaram-se três dados na safra de 2003/04 em 40 e 200 kg ha⁻¹ ano⁻¹ de N e na safra de 2004/05 no tratamento 40 kg.ha⁻¹ ano⁻¹ de N. Os elementos B e Mn apresentaram apenas uma observação (Figura 7). discrepantes na safra de 2005/06, sendo o B nas doses 80 e 200 kg ha⁻¹ ano⁻¹ de N e Mn em 40 kg ha⁻¹ ano⁻¹ de N (Figura 7).

Campo Largo. Do universo de 60 dados, foram constatados pontos *outliers* para o elemento P nas doses 120 e 160 kg.ha⁻¹ ano⁻¹ de N, na safra de 2003/04, o mesmo ocorreu no tratamento 120 kg.ha⁻¹ ano⁻¹ de N na safra de 2004/05. O K apresentou um ponto distante da média em 200 kg.ha⁻¹ ano⁻¹ de N, na safra de 2003/04 (Figura 8).

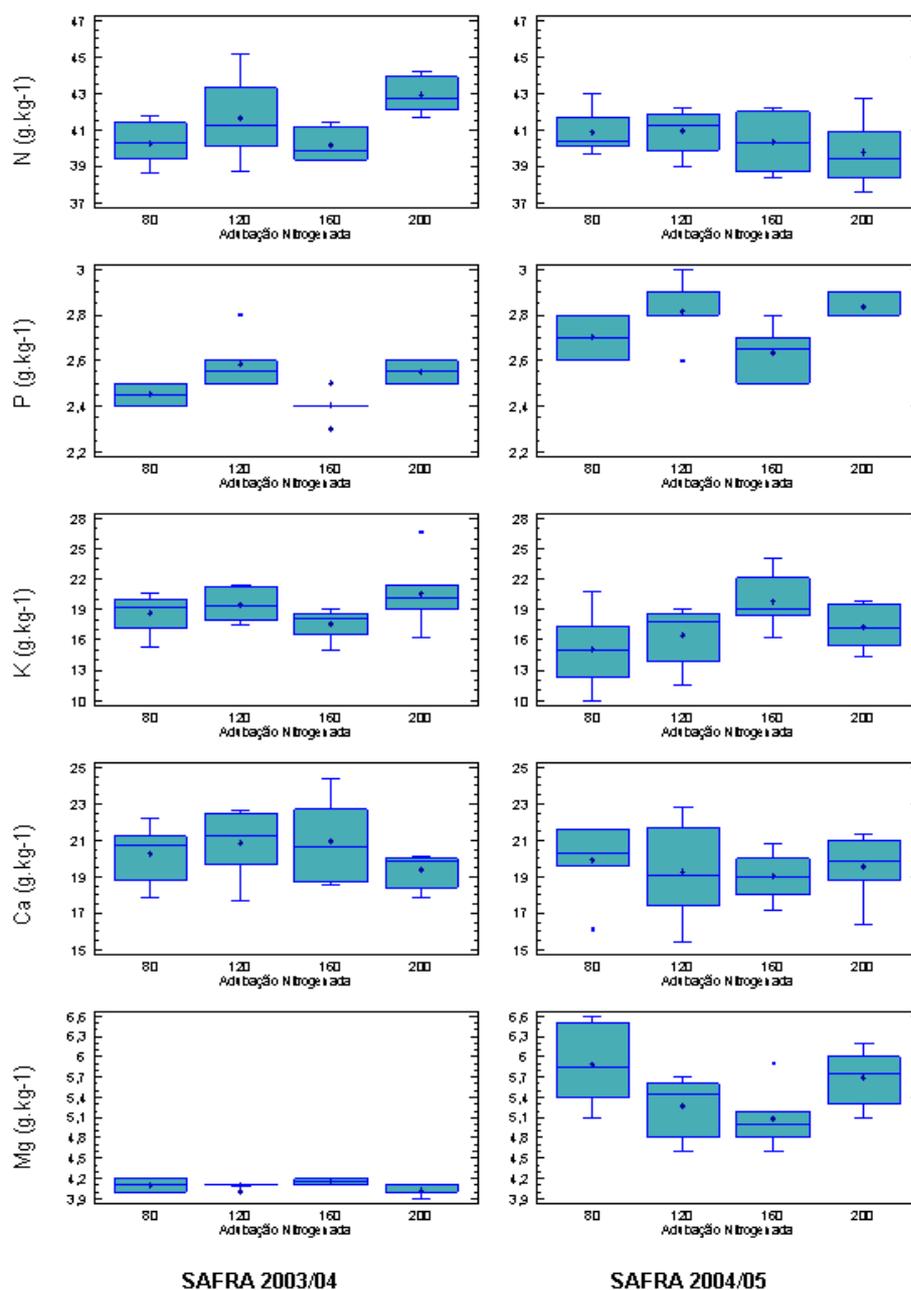


FIGURA 9 - *BOX-PLLOT* DOS MACRONUTRIENTES, NOS DIFERENTES DOSES DE ADUBAÇÕES NITROGENADAS ($\text{kg ha}^{-1} \text{ano}^{-1}$) REALIZADAS NO POMAR DE PÊSSEGO 'CHIMARRITA' EM CAMPO LARGO, PR.

Para o Ca em 80 kg.ha⁻¹ ano⁻¹ de N na safra 2004/05 e para o Mg é possível observar um ponto no tratamento 200 kg.ha⁻¹ ano⁻¹ de N na safra 2004/05. As demais informações sobre os macronutrientes não apresentam problemas de dispersão de dados em torno da média de cada tratamento (Figura 8).

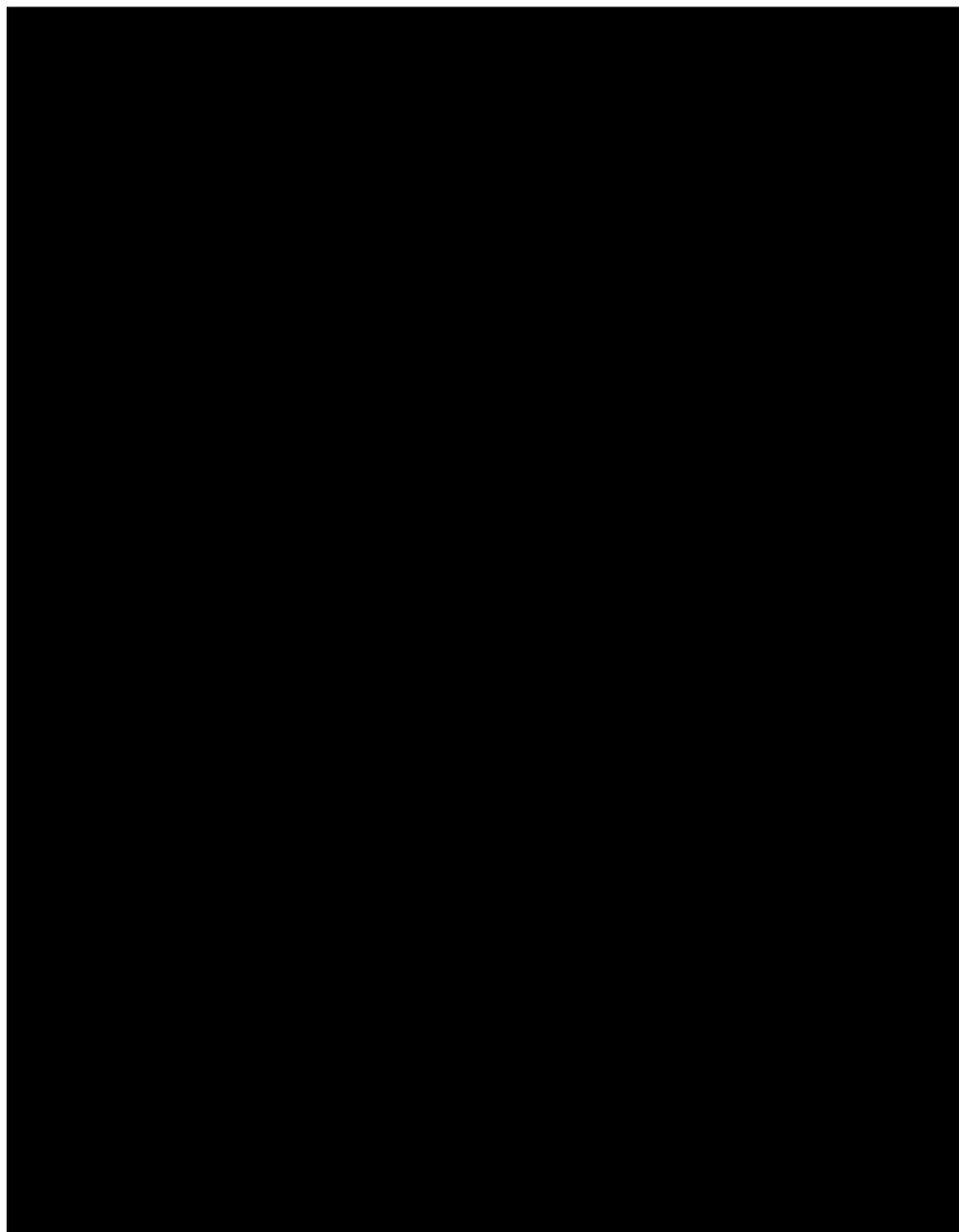


FIGURA 10 - BOX PLOT DOS MICRONUTRIENTES, NOS DIFERENTES DOSES DE ADUBAÇÕES NITROGENADAS (kg ha⁻¹ ano⁻¹) REALIZADAS NO POMAR DE PÊSSEGO 'CHIMARRITA' EM CAMPO LARGO, PR.

Na figura 9 pode-se verificar que o que Fe na presença de 80 kg.ha⁻¹ ano⁻¹ de N, em ambas as safras e 160 kg.ha⁻¹ ano⁻¹ de N na safra de 2003/04, apresentam pontos *outliers*.

O mesmo ocorre com Cu nas doses de 80 e 120 kg.ha⁻¹ ano⁻¹ de N na safra de 2004/05. O Zn apresentou um ponto distante da média em 80 kg.ha⁻¹ ano⁻¹ de N, na safra de 2004/05.

O B apresentou pontos discrepantes em 80 e 160 kg.ha⁻¹ ano⁻¹ de N na safra de 2003/04 e no tratamento 120 kg.ha⁻¹ ano⁻¹ de N na safra de 2004/05. O Mn apresentou somente um ponto *outlier* na adubação 160 kg.ha⁻¹ ano⁻¹ de N. As demais variáveis não apresentaram dados discrepantes (Figura 9).

Os pontos discrepantes apresentados nos *Box-plot* (Figura 6, 7, 8 e 9), quando comparados com os níveis teóricos constantes no Manual de Adubação e Calagem (2004), se apresentam dentro das faixas de interpretação (Figura 2, 3, 4 e 5) e com a dispersão corroborando com a prática. Portanto, os dados se encontram dentro dos padrões práticos interpretados a campo.

O método de apresentar os dados por meio do gráfico *Box-plot* possibilita comparar a dispersão dos dados e ao mesmo tempo, a diferença entre os fatores, bem como, identifica pontos discrepantes. Desta forma esta técnica aborda, de maneira robusta, as conclusões obtidas por meio das análises dos CV's e de dispersão apresentadas anteriormente. Outro, porém, pode ser usado como um indicativo para prever resultados obtidos por meio da ANOVA.

Segundo estudos realizados por ALVES e CUNHA (2006), o *software* Statgraphics utilizado nessas análises, é apontado como fornecedor de excelentes desenhos esquemáticos (*Box-plot*), entretanto, ao empregar este gráfico, deve se precaver ao utilizar versões diferentes ou outros programas variados na mesma base de dados, pois questões de precisão nos cálculos estatísticos e diferenças de definição podem gerar resultados ligeiramente diversos (BENJAMINI, 1988). Também deve ser mantida constante a precisão nas investigações, evitando trocas do programa.

Verifica-se que a análise descritiva, realizada nos dois pomares, facilita a interpretação dos resultados, descrevendo-os resumidamente com suas principais estatísticas (TOLEDO e OVALLE, 1995), identificando características que possam evidenciar futuros problemas na validação dos pressupostos da ANOVA. Neste caso, sugere-se o uso de outra metodologia de análise estatística, que não seja fundamentada pela distribuição normal e/ou pela homocedasticidade dos resíduos.

A metodologia sugerida por PIMENTEL GOMES (1963) para experimentos com mais de duas séries de parâmetros, analisa o relacionamento entre as variáveis duas a duas, o qual recorre ao estudo da regressão ou da correlação entre as variáveis (COSTA, 1977).

Por meio da Análise de Correlação de Spearman e Pearson verifica-se a relação dos teores foliares NPK nos pomares da Lapa e Campo Largo (CAVADIAS, PILON e YUE, 2002).

3.4 ANÁLISE DE CORRELAÇÃO

Lapa. Observa-se nas safras de 2003/04 e 2004/05 uma relação moderada positiva ($0,68 < \text{Coef.}_{(p\text{-valor} < 0,001)} < 0,70$) entre as diferentes doses de adubação nitrogenada e a quantidade de N encontrado na folha.

O N tem correlação inversa com a média de N encontrado nos blocos ($\text{Coef.}_{(p\text{-valor} < 0,01)} = 0,42$), correlação positiva com o P nas três safras ($0,41 < \text{Coef.}_{(p\text{-valor} < 0,05)} < 0,56$) e na safra de 2005/06 com K ($\text{Coef.}_{(p\text{-valor} < 0,001)} < 0,48$). Com relação à variável P é possível identificar uma correlação com o K na safra de 2004/05 ($\text{Coef.}_{(p\text{-valor} < 0,001)} < 0,46$). Não foi possível identificar correlação com as doses da adubação potássica com as demais variáveis (Figura 10).

Analisando os elementos NPK em relação aos anos, pode-se observar uma tendência em diminuir o N e o P e aumentar o K, sendo isso, evidenciado e confirmado pela análise de correlação de Spearman. Portanto, estatisticamente NP teve uma perda ao longo do tempo observado e um ganho de K (Figura 4 e 10).

Pode ser visualizado na Tabela 9, o coeficiente de correlação significativo nas safras de 2003/04 e 2004/0 no fator bloco, afirmando a necessidade da utilização de blocos no experimento.

Corroborando com análise por meio do *Box-plot* (Figura 6) é verificado a ausência de associação entre adubação N e N na planta no ano 2005/06.

	kg.ha ⁻¹ ano ⁻¹			Unit.	g.kg ⁻¹			kg.ha ⁻¹ ano ⁻¹			Unit.	g.kg ⁻¹			Anos
	N	K ₂ O	BLOCO		N	P	K	N	K ₂ O	BLOCO		N	P	K	Safra
SAFRA 03/2004	N (kg.ha ⁻¹ ano ⁻¹)	1,00	0,00	0,00	0,68	0,32	-0,28	1,00	0,00	0,00	0,70	0,56	-0,12		
	p-valor	,	1,00	1,00	0,00	0,08	0,13	,	1,00	1,00	0,00	0,00	0,52		
	K ₂ O (kg.ha ⁻¹ ano ⁻¹)	0,00	1,00	0,00	0,05	0,00	-0,06	0,00	1,00	0,00	-0,13	-0,06	0,03		
	p-valor	1,00	,	1,00	0,79	1,00	0,75	1,00	,	1,00	0,50	0,76	0,87		
	BLOCO (Unit.)	0,00	0,00	1,00	-0,42	-0,33	-0,15	0,00	0,00	1,00	-0,30	0,05	0,53		
	p-valor	1,00	1,00	,	0,02	0,08	0,42	1,00	1,00	,	0,11	0,79	0,00		
	N (g.kg ⁻¹)	0,68	0,05	-0,42	1,00	0,56	-0,18	0,70	-0,13	-0,30	1,00	0,54	-0,15		
	p-valor	0,00	0,79	0,02	,	0,00	0,34	0,00	0,50	0,11	,	0,00	0,43		
	P (g.kg ⁻¹)	0,32	0,00	-0,33	0,56	1,00	-0,36	0,56	-0,06	0,05	0,54	1,00	0,46		
	p-valor	0,08	1,00	0,08	0,00	,	0,05	0,00	0,76	0,79	0,00	,	0,01		
	K (g.kg ⁻¹)	-0,28	-0,06	-0,15	-0,18	-0,36	1,00	-0,12	0,03	0,53	-0,15	0,46	1,00		
	p-valor	0,13	0,75	0,42	0,34	0,05	,	0,52	0,87	0,00	0,43	0,01	,		
SAFRA 04/2005	N (kg.ha ⁻¹ ano ⁻¹)	1,00	0,00	0,00	0,34	-0,05	0,30	1,00	0,00	0,00	0,32	0,30	0,04	0,00	
	p-valor	,	1,00	1,00	0,06	0,79	0,11	,	1,00	1,00	0,00	0,00	0,69	1,00	
	K ₂ O (kg.ha ⁻¹ ano ⁻¹)	0,00	1,00	0,00	0,12	0,12	0,19	0,00	1,00	0,00	-0,02	0,05	0,05	0,00	
	p-valor	1,00	,	1,00	0,54	0,53	0,33	1,00	,	1,00	0,87	0,66	0,66	1,00	
	BLOCO (Unit.)	0,00	0,00	1,00	0,22	0,66	0,12	0,00	0,00	1,00	-0,13	0,07	0,13	0,00	
	p-valor	1,00	1,00	,	0,23	0,00	0,54	1,00	1,00	,	0,22	0,50	0,23	1,00	
	N (g.kg ⁻¹)	0,34	0,12	0,22	1,00	0,41	0,48	0,32	-0,02	-0,13	1,00	0,55	-0,47	-0,73	
	p-valor	0,06	0,54	0,23	,	0,02	0,01	0,00	0,87	0,22	,	0,00	0,00	0,00	
	P (g.kg ⁻¹)	-0,05	0,12	0,66	0,41	1,00	0,10	0,30	0,05	0,07	0,55	1,00	-0,24	-0,39	
	p-valor	0,79	0,53	0,00	0,02	,	0,60	0,00	0,66	0,50	0,00	,	0,02	0,00	
	K (g.kg ⁻¹)	0,30	0,19	0,12	0,48	0,10	1,00	0,04	0,05	0,13	-0,47	-0,24	1,00	0,68	
	p-valor	0,11	0,33	0,54	0,01	0,60	,	0,69	0,66	0,23	0,00	0,02	,	0,00	
Safra (Anos)							0,00	0,00	0,00	-0,73	-0,39	0,68	1,00		
p-valor							1,00	1,00	1,00	0,00	0,00	0,00	,		
SAFRA 05/2006	N (kg.ha ⁻¹ ano ⁻¹)	1,00	0,00	0,00	0,34	-0,05	0,30	1,00	0,00	0,00	0,32	0,30	0,04	0,00	
	p-valor	,	1,00	1,00	0,06	0,79	0,11	,	1,00	1,00	0,00	0,00	0,69	1,00	
	K ₂ O (kg.ha ⁻¹ ano ⁻¹)	0,00	1,00	0,00	0,12	0,12	0,19	0,00	1,00	0,00	-0,02	0,05	0,05	0,00	
	p-valor	1,00	,	1,00	0,54	0,53	0,33	1,00	,	1,00	0,87	0,66	0,66	1,00	
	BLOCO (Unit.)	0,00	0,00	1,00	0,22	0,66	0,12	0,00	0,00	1,00	-0,13	0,07	0,13	0,00	
	p-valor	1,00	1,00	,	0,23	0,00	0,54	1,00	1,00	,	0,22	0,50	0,23	1,00	
	N (g.kg ⁻¹)	0,34	0,12	0,22	1,00	0,41	0,48	0,32	-0,02	-0,13	1,00	0,55	-0,47	-0,73	
	p-valor	0,06	0,54	0,23	,	0,02	0,01	0,00	0,87	0,22	,	0,00	0,00	0,00	
	P (g.kg ⁻¹)	-0,05	0,12	0,66	0,41	1,00	0,10	0,30	0,05	0,07	0,55	1,00	-0,24	-0,39	
	p-valor	0,79	0,53	0,00	0,02	,	0,60	0,00	0,66	0,50	0,00	,	0,02	0,00	
	K (g.kg ⁻¹)	0,30	0,19	0,12	0,48	0,10	1,00	0,04	0,05	0,13	-0,47	-0,24	1,00	0,68	
	p-valor	0,11	0,33	0,54	0,01	0,60	,	0,69	0,66	0,23	0,00	0,02	,	0,00	
Safra (Anos)							0,00	0,00	0,00	-0,73	-0,39	0,68	1,00		
p-valor							1,00	1,00	1,00	0,00	0,00	0,00	,		

* Análise de Correlação de Spearman nas variáveis relacionadas com os tratamentos (N, K₂O e Bloco).

** Análise de Correlação de Pearson entre as variáveis NPK.

FIGURA 11 - ANÁLISE DE CORRELAÇÃO DE NPK, SAFRAS 2003/04, 2004/05 E 2005/06, NA PRESENÇA DE DIFERENTES DOSES DE ADUBAÇÕES NITROGENADAS E POTÁSSICA (kg ha⁻¹ ano⁻¹) DOS POMARES DE PÊSSEGO 'CHIMARRITA'. LAPA, PR.

Campo Largo. Na figura 11, é possível verificar uma associação moderada positiva (Coef._(p-valor < 0,001) = 0,67) entre a média de K encontrado no bloco I com K encontrado na folha, portanto, à medida que se parte do bloco I passando pelo bloco 2 e finalizando no bloco III aumenta a quantidade de K na safra de 2003/04. Também é possível verificar uma relação positiva e moderada entre os teores foliares N e P (Coef._(p-valor < 0,001) = 0,67). Nas demais variáveis apresentadas na Figura 11, safra de 2003/04 não é possível afirmar estatisticamente associações. Na safra de 2004/05, a única associação identificada foi entre K e bloco, (Coef._(p-valor = 0,04) = 0,44) moderada positiva.

	kg.ha ⁻¹ ano ⁻¹			unit.	g.kg ⁻¹			kg.ha ⁻¹ ano ⁻¹			unit.	g.kg ⁻¹			ano
	N	K ₂ O	BLOCO	N	P	K	N	K ₂ O	BLOCO	N	P	K	Safra		
SAFRA 03/2004	N kg.ha ⁻¹ ano ⁻¹	1,00	0,00	0,00	0,40	0,20	0,11	1,00	0,00	0,00	0,07	0,12	0,19	0,00	
	p-valor	,	1,00	1,00	0,05	0,34	0,63	,	1,00	1,00	0,66	0,42	0,20	1,00	
	K ₂ O kg.ha ⁻¹ ano ⁻¹	0,00	1,00	0,00	0,10	0,04	0,07	0,00	1,00	0,00	-0,16	-0,01	0,16	0,00	
	p-valor	1,00	,	1,00	0,63	0,84	0,76	1,00	,	1,00	0,28	0,96	0,27	1,00	
	BLOCO unit.	0,00	0,00	1,00	-0,41	-0,04	0,67	0,00	0,00	1,00	-0,32	0,06	0,53	0,00	
	p-valor	1,00	1,00	,	0,05	0,84	0,00	1,00	1,00	,	0,03	0,67	0,00	1,00	
	N g.kg ⁻¹	0,40	0,10	-0,41	1,00	0,67	-0,05	0,07	-0,16	-0,32	1,00	0,06	-0,07	-0,19	
	p-valor	0,05	0,63	0,05	,	0,00	0,82	0,66	0,28	0,03	,	0,69	0,66	0,19	
	P g.kg ⁻¹	0,20	0,04	-0,04	0,67	1,00	0,26	0,12	-0,01	0,06	0,06	1,00	-0,12	0,75	
	p-valor	0,34	0,84	0,84	0,00	,	0,22	0,42	0,96	0,67	0,69	,	0,40	0,00	
	K g.kg ⁻¹	0,11	0,07	0,67	-0,05	0,26	1,00	0,19	0,16	0,53	-0,07	-0,12	1,00	-0,30	
	p-valor	0,63	0,76	0,00	0,82	0,22	,	0,20	0,27	0,00	0,66	0,40	,	0,04	
	Safra ano							0,00	0,00	0,00	-0,19	0,75	-0,30	1,00	
	p-valor							1,00	1,00	1,00	0,19	0,00	0,04	,	
SAFRA 04/2005	N kg.ha ⁻¹ ano ⁻¹	1,00	0,00	0,00	-0,29	0,25	0,31								
	p-valor	,	1,00	1,00	0,17	0,23	0,14								
	K ₂ O kg.ha ⁻¹ ano ⁻¹	0,00	1,00	0,00	-0,43	-0,05	0,27								
	p-valor	1,00	,	1,00	0,04	0,82	0,20								
	BLOCO unit.	0,00	0,00	1,00	-0,24	0,33	0,44								
	p-valor	1,00	1,00	,	0,25	0,12	0,03								
	N g.kg ⁻¹	-0,29	-0,43	-0,24	1,00	0,05	-0,25								
	p-valor	0,17	0,04	0,25	,	0,83	0,24								
	P g.kg ⁻¹	0,25	0,05	0,33	0,05	1,00	0,12								
	p-valor	0,23	0,82	0,12	0,83	,	0,56								
K g.kg ⁻¹	0,31	0,27	0,44	-0,25	0,12	1,00									
p-valor	0,14	0,20	0,03	0,24	0,56	,									

*Análise de Correlação de Spearman nas variáveis relacionadas com os tratamentos (N, K₂O e Bloco).

**Análise de Correlação de Pearson entre as variáveis NPK.

FIGURA 12 - ANÁLISE DE CORRELAÇÃO DE NPK, SAFRAS 2003/04 E 2004/05, NA PRESENÇA DE DIFERENTES DOSES DE ADUBAÇÕES NITROGENADAS E POTÁSSICA (kg ha⁻¹)

Nos estudos realizados por TREVISAN, et. al., (2006), utilizaram o teste de Correlação de Pearson para verificar a relação entre sabor estranho e odor estranho e amargo e odor estranho sensorial apresentando uma correlação moderada ($R^2 = 0,86$ e $0,64$ respectivamente). No entanto, neste trabalho não foi considerada a característica qualitativa da variável resposta, bem como, não foi apresentado o p-valor do teste de significância, resultados que podem validar ou invalidar as conclusões. O mesmo foi observado no trabalho realizado por SCHUHMACHER em que avaliou por meio da análise de correlação a associação entre os teores foliares do milho (*Zea mays L.*)

Quando o interesse é verificar a relação entre conjuntos de variáveis explicativas e resposta a metodologia indicada é Análise de Correlação Canônica. Esta pode ser vista como uma extensão da regressão múltipla que busca a máxima correlação entre variáveis independentes e dependentes (HAIR, et. al., 1995).

3.5 ANÁLISE DE CORRELAÇÃO CANÔNICA

Lapa. A Figura 12 deslumbra as correlações multivariadas entre as três safras. Desta forma é possível afirmar por meio da correlação canônica que os teores foliares de macro e micronutrientes da safra de 2003/04 apresentaram alta associação positiva com as safras 2004/05 (Coef. $(p\text{-valor} < 0,002) = 0,97$) e 2005/06 (Coef. $(p\text{-valor} < 0,001) = 0,96$). O mesmo observou-se entre as safras de 2004/05 e 2005/06 (Coef. $(p\text{-valor} < 0,01) = 0,91$). Portanto, os macronutrientes e micronutrientes estão altamente correlacionados entre as três safras observadas. O teor foliar entre os macros e micronutrientes permanecem proporcionais entre si durante todas as safras.

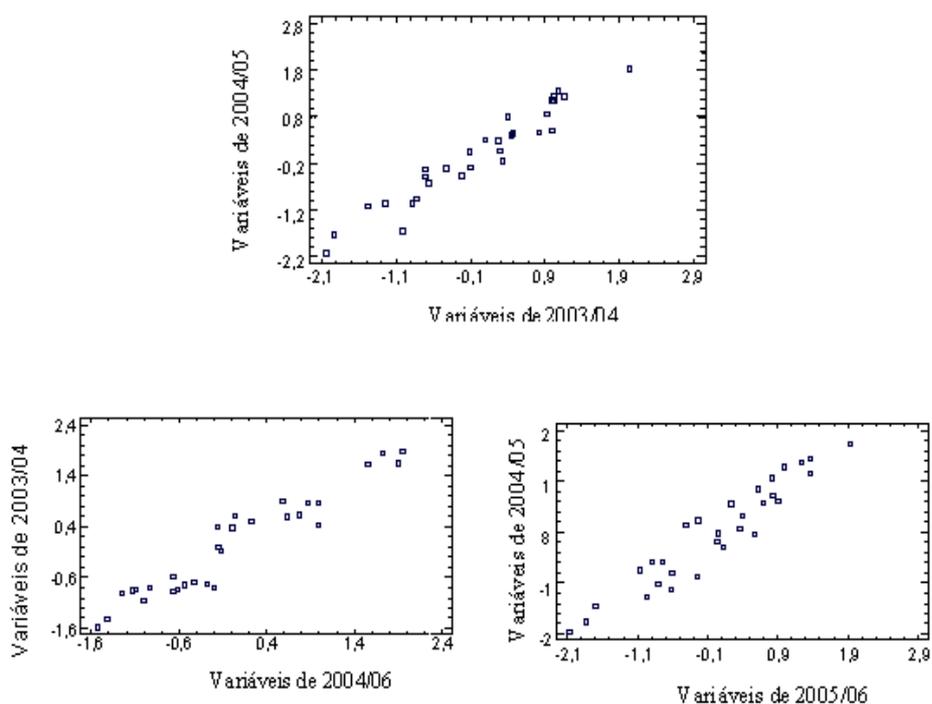


FIGURA 13 - ANÁLISE DE CORRELAÇÃO CANÔNICA DOS TEORES FOLIARES DE MACRO E MICRO NUTRIENTES DAS SAFRAS DE 2003/04, 2004/05, 2005/06 EM POMAR DE PÊSSEGO 'CHIMARRITA'. LAPA, PR.

Entretanto, ao considerar somente as variáveis NPK, há uma correlação entre as safras 2003/04 e 2004/05 (Coef. (p-valor < 0,01) = 0,78). Nas demais comparações o poder do teste de correlação não foi significativo, apresentando os seguintes resultados: entre as safras 2004/05 e 2005/06 (Coef. (p-valor = 0,0755) = 0,5826) e entre as safras 2003/04 e 2005/06 (Coef.(p-valor = 0,4803) = 0,3957). Partindo dos resultados obtidos na Figura 6, é possível que outras variáveis estejam influenciando em NPK, dentre as safras.

Campo Largo. A Figura 13 apresenta a correlação canônica significativamente positiva das diagnoses foliares entre as safras observadas. Os teores foliares da safra de 2003/04 apresentaram alta associação positiva com as safras 2004/05 (Coef. (p-valor < 0,05) = 0,97). Portanto, os macronutrientes e micronutrientes estão altamente correlacionados entres as duas safras observadas.

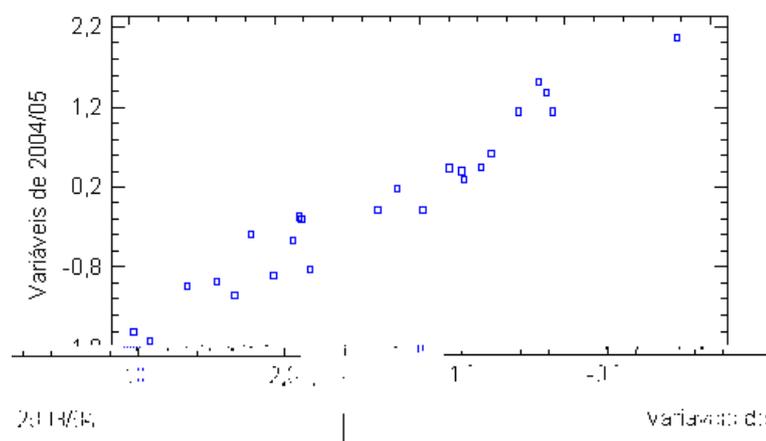


FIGURA 14 - ANÁLISE DE CORRELAÇÃO CANÔNICA DA DIAGNOSE FOLIAR DAS SAFRAS DE 2003/04 E 2004/05 EM POMAR DE PÊSSEGO 'CHIMARRITA'. CAMPO LARGO, PR.

Quando comparado os métodos de análises de correlações: a vantagem da análise de correlação canônica sobre as demais é que possibilita quantificar associação entre dois conjuntos de variáveis. O que não é possível ser realizado com análise de correlação de Spearman e Pearson ou pelo gráfico de *Box-plot*.

Análise de correlação de Spearman é eficiente quando objetivo é quantificar o grau de associação entre variáveis categóricas e variáveis contínuas, ou seja, relação entre doses de nitrogênio com teor de N encontrado na folha do pessegueiro (CAVADIAS, PILON e YUE, 2002).

Quando em um experimento trabalha-se apenas com uma única série de variáveis (parâmetros) o procedimento correto é a análise da variância (ANOVA) e o teste entre médias (PIMENTEL GOMES, 1963). Portanto, outra forma de analisar os dados é por meio da ANOVA, que consiste em avaliar simultaneamente os tratamentos realizados na planta, testando a existência de diferenças significativas entre os contrastes das médias, decompondo a variação total das observações em partes conhecidas e em partes desconhecidas e não controláveis (resíduos ou erro) (MONTGOMERY, 2001).

As causas desconhecidas muitas vezes inviabilizam análise, chegando a perder o experimento realizado.

Depois de verificar os pressupostos de a ANOVA a seguir, abordou os teores foliares de NPK de ambos os pomares, decompondo a variação total das observações, segundo o ano, adubação nitrogenada e potássica.

3.6 ANÁLISE DE VARIÂNCIA

Verifica-se na Tabela 3 no pomar da Lapa, que a quantidade de teores foliares N e K de um ano, difere dos demais anos (Teste $F_{p\text{-valor}} < 0,05$). Também se observa que, pelo menos uma das doses de nitrogênio aplicado no solo, difere das demais, ao considerar como variável resposta a quantidade de N. A variável P não apresentou fatores significantes (Teste $F_{p\text{-valor}} > 0,058$), também verificado graficamente nas análises descritivas. No mesmo sentido, a blocagem não foi significativa (Teste $F_{p\text{-valor}} > 0,167$) nas variáveis NPK (Tabela 3), portanto não haveria a necessidade de utilização de blocos, discordando da conclusão obtida na análise de correlação. Entretanto ANOVA é o método mais indicado para afirmar a necessidade da presença de blocagem no experimento.

No pomar de Campo Largo os resultados foram diferentes aos encontrados no pomar da Lapa, isso pode ser explicado pela diferença no manejo entre os pomares. Como a necessidade da utilização de blocos no experimento quando a variável de interesse é K (Teste $F_{p\text{-valor}} < 0,004$). Outro, porém, a quantidade do P e K encontrado na folha depende do ano (Teste $F_{0,005 < p\text{-valor} < 0,020}$). Já os teores foliar de N e K nas doses de nitrogênio aplicada no solo depende do ano (Teste $F_{0,003 < p\text{-valor} < 0,0381}$).

A variável P apresentou contrastes significativos nas doses de N (Teste $F_{p\text{-valor}} < 0,001$) (Tabela 3).

Os pontos discrepantes identificados na análise descritiva não influenciaram nos pressupostos exigidos pela ANOVA de ambos os pomares.

Com base nos resultados apresentados na Tabela 3, é possível verificar fatores significativos, portanto, para avaliar a magnitude das diferenças encontradas na ANOVA, se faz necessário realizar teste de comparações múltiplas, permitindo testar qualquer contraste. Na Tabela 2 é apresentado os resultados das comparações múltiplas das médias obtidas por meio do Teste de Tukey a 5% de significância.

TABELA 5 - ANÁLISE DE VARIÂNCIA DOS RESULTADOS DE NITROGENIO (N g.kg⁻¹), FÓSFORO (P g.kg⁻¹) E POTÁSSIO (K g.kg⁻¹) ENCONTRADO NAS FOLHAS, EM PRESENÇA DE DIFERENTES DOSES DE ADUBAÇÃO NITROGENADA E POTÁSSICA, NO POMAR DE PÊSSEGO 'CHIMARRITA', LOCALIZADO NA LAPA E CAMPO LARGO, PR.

VARIÁVEIS		NITROGÊNIO	FÓSFORO	POTÁSSIO
C.V	GL	p-valor		
POMAR DA LAPA				
BLOCO	2	0,167	0,846	0,467
ANO	2	0,001	0,194	0,013
ERRO (BLOCO)	4			
PARCELA				
ADUBAÇÃO N	4	0,000	0,058	0,313
ANO x ADUBAÇÃO N	8	0,224	0,615	0,769
ERRO (ANO)	24			
SUBPARCELA				
ADUBAÇÃO K	1	0,678	0,547	0,470
ANO x ADUBAÇÃO K	2	0,636	0,741	0,256
ADUBAÇÃO N x ADUBAÇÃO K	4	0,658	0,417	0,273
ANO x N x K	8	0,187	0,381	0,441
ERRO (N)	30			
TOTAL	89			
POMAR DE CAMPO LARGO				
BLOCO	2	0,102	0,636	0,004
ANO	1	0,112	0,020	0,005
ERRO (BLOCO)	2			
PARCELA				
ADUBAÇÃO N	3	0,052	0,001	0,200
ANO x ADUBAÇÃO N	3	0,003	0,911	0,038
ERRO (ANO)	12			
SUBPARCELA				
ADUBAÇÃO K	1	0,252	1,000	0,091
ANO x ADUBAÇÃO K	1	0,009	0,248	0,444
ADUBAÇÃO N x ADUBAÇÃO K	3	0,125	0,047	0,522
ANO x N x K	3	0,009	0,195	0,206
ERRO (N)	16			
TOTAL	47			

* Nível de significância $\leq 0,05$.

Na elaboração de diagnósticos nutricionais, usualmente comparam-se os resultados de análise foliar com valores padrão (doses críticas ou faixas de suficiências), admitindo-se que nutrientes inferiores a esses, provavelmente estejam limitando o crescimento, rendimento ou

qualidade dos frutos (MARTIN-PRÉVEL, 1990; LAHAV, 1995), portanto, por meio do Manual de adubação e calagem para os Estados do Rio Grande do Sul e Santa Catarina verificaram-se os padrões dos teores foliares (Figura 2, 3, 4 e 5). No trabalho sobre efeitos da adubação NPK na fertilidade do solo, nutrição e crescimento da seringueira realizado por BATAGLIA e SANTOS (1999), mostrou que adubação nitrogenada (0, 40, 80 e 120 kg.ha⁻¹) não influenciou na concentração de NPK encontrado na folha. Outro, porém, a concentração encontrada de NPK na folha foi semelhante aos encontrados nos pomares da Lapa e Campo Largo. Por outro lado, os resultados significativos dos teores de N em relação às datas de amostragem corroboram com os resultados encontrados por apresentado BATAGLIA e SANTOS (1999) em que realizou amostragem aos 60 e 72 meses na planta a qual tem uma relação moderada com teor de Nitrogênio encontrado na folha. Também é possível verificar resultados semelhantes no trabalho de DOLINSKI et al. (2005), citado por SOUZA (2005) em que compara o teor de N na folha do pessegueiro entre três safras (2002/03, 2003/04 e 2004/05), mostrando que a concentração aumentou apenas na safra de 2004/05.

A ANOVA é um meio de quantificar exatamente a variabilidade dentro de cada fator de interesse, o que não é possível verificar pelas figuras esquemáticas *Box-plot*.

Mas como o Box-plot e análise de dispersão, a ANOVA também pode ser usado como um indicativo da existência ou não de diferenças significativas entre os fatores de interesse, com a vantagem de ter maior precisão. Entretanto, por ser mais eficiente há necessidade da verificação dos pressupostos de normalidade, homocedasticidade, independência e aditividade dos resíduos gerados pelo modelo matemático gerado por meio da ANOVA.

Para afirmar diferenças significativas em pelos menos um dos níveis de cada fator em relação aos demais níveis, é necessária a realização da análise de comparação múltipla. Portanto, a ANOVA passa a ter somente a função de apresentar os resultados referentes ao comportamento dos resíduos do modelo.

3.7 ANÁLISE DE COMPARAÇÕES MÚLTIPLAS

Lapa. Verifica-se o teor foliar de N reduziu 19% da safra de 2003/04 para as safras de 2004/05 e 2005/06. Entretanto, o teor foliar de K aumentou 13% na safra de 2005/06 em relação às safras 2003/04 e 2004/05. Avaliando individualmente as safras, verifica-se que os teores foliares de N na presença de mais de 120 kg.ha⁻¹ N, aumentaram 13% e 19% nas safras 2003/04 e 2004/05 respectivamente. A adubação nitrogenada e potássica não influenciaram o

teor de N na safra de 2005/06, corroborando com análise de correlação, onde apenas nos dois primeiros anos verifica-se o efeito da aplicação de N.

Os teores foliares de N na safra 2003/04, na presença das doses 80 a 200 kg.ha⁻¹ ano⁻¹ de N e 55 a 110 kg.ha⁻¹ ano⁻¹ de K, encontravam-se normais, entretanto, os demais resultados estavam abaixo do normal (Tabela 4, Figura 2).

O teor foliar de P encontrou-se em níveis normais durante as três safras, independente das adubações e safras (Tabela 4, Figura 2), embora análise de correlação tenha mostrado associação entre aplicação de N sobre P..

Em todas as safras o K encontrado na folha não sofreu influência das adubações, aumentando 13% entre as safras 2003/04 e 2005/06. O teor foliar de K se encontrou em níveis acima do normal (Tabela 4, Figura 2).

TABELA 6 - ANÁLISE DOS RESULTADOS DE NITROGENIO (N g.kg⁻¹), FÓSFORO (P g.kg⁻¹) E POTÁSSIO (K g.kg⁻¹) ENCONTRADOS NA FOLHA, EM PRESENÇA DE DIFERENTES DOSES DE ADUBAÇÃO NITROGENADA E POTÁSSICA, NO POMAR DE PÊSSEGO 'CHIMARRITA', LOCALIZADO NA LAPA, PR.

Tratamentos kg.ha ⁻¹ .ano ⁻¹	Safras		
	2003/04	2004/05	2005/06
Nitrogênio	N (g.kg⁻¹)		
40	31,5 Ab	24,9 Aa	26,1 Aa
80	33,9 Ba	26,6 ABb	27,1 Ab
120	34,4 BCa	28,8 BCb	26,8 Ab
160	36,1 Ca	30,7 Cb	28,7 Ab
200	36,2 Ca	29,5 Cb	26,8 Ab
Potássio			
55	34,4 Aa	28,4 Aa	27,0 Ab
110	34,5 Aa	27,8 Ab	27,1 Ab
Nitrogênio	P (g.kg⁻¹)		
40	2,28 Aa	2,21 Aa	2,25 Aa
80	2,33 Aa	2,28 Aa	2,25 Aa
120	2,32 Aa	2,33 Aa	2,23 Aa
160	2,40 Aa	2,35 Aa	2,28 Aa
200	2,37 Aa	2,37 Aa	2,23 Aa
Potássio			
55	2,33 Aa	2,31 Aa	2,24 Aa
110	2,34 Aa	2,30 Aa	2,26 Aa
Nitrogênio	K (g.kg⁻¹)		
40	20,7 Aa	21,8 Aab	22,7 Ab
80	20,7 Aa	22,3 Ab	23,6 Ac
120	21,3 Aa	22,5 Aa	23,0 Aa
160	19,9 Aa	21,3 Aa	23,4 Ab
200	20,7 Aa	22,1 Aab	24,1 Ab
Potássio			
55	20,8 Aa	21,9 Aab	23,1 Ab
110	20,5 Aa	22,1 Ab	23,6 Ac

Teste Tukey, letras iguais minúsculas na linha e maiúsculas na coluna não diferem estatisticamente com nível de 5% significância.

Análise química realizada pela EMBRAPA - Clima temperado – Pelotas – RS, Laboratório de Nutrição Mineral de Plantas.

Campo Largo. Na safra de 2003/04, o teor foliar de N, na presença da adubação $160 \text{ kg.ha}^{-1} \text{ ano}^{-1}$ de N aumentou 7%, não diferindo da concentração de $200 \text{ kg.ha}^{-1} \text{ ano}^{-1}$ de N. Na safra de 2004/05 as diferentes doses de adubação nitrogenada não influenciaram na quantidade do teor foliar de N. Entre as safras de 2003/04 e 2004/05, os tratamentos que apresentaram diferenças significativas de um ano para outro foram as doses de $200 \text{ kg.ha}^{-1} \text{ ano}^{-1}$ de N e $110 \text{ kg.ha}^{-1} \text{ ano}^{-1}$ de K, apresentando uma diminuição na quantidade de 7% e 4% respectivamente. Em ambas as safras o teor foliar de N se manteve normal (Tabela 5, Figura 4).

Geralmente quando mais de dois fatores interage, análise dos dados se torna complexa baseada somente na tabela de comparação de médias, conforme descrito anteriormente. A forma mais prática de verificar as interações é graficamente.

Desta forma, os resultados verificados para a concentração de N na folha e baseado na significância do teste F (Tabela 3) com relação à interação entre adubação nitrogenada, potássica e as safras, pode ser facilmente visualizado na Figura 14.

O cruzamento das linhas é um indicativo que os fatores estão interagindo.

Verifica-se que o teor de N foi menor na safra de 2004/05 quando verificada na presença da dose $200 \text{ kg.ha}^{-1} \text{ ano}^{-1}$ de N e $110 \text{ kg.ha}^{-1} \text{ ano}^{-1}$ de K (Figura 14 A).

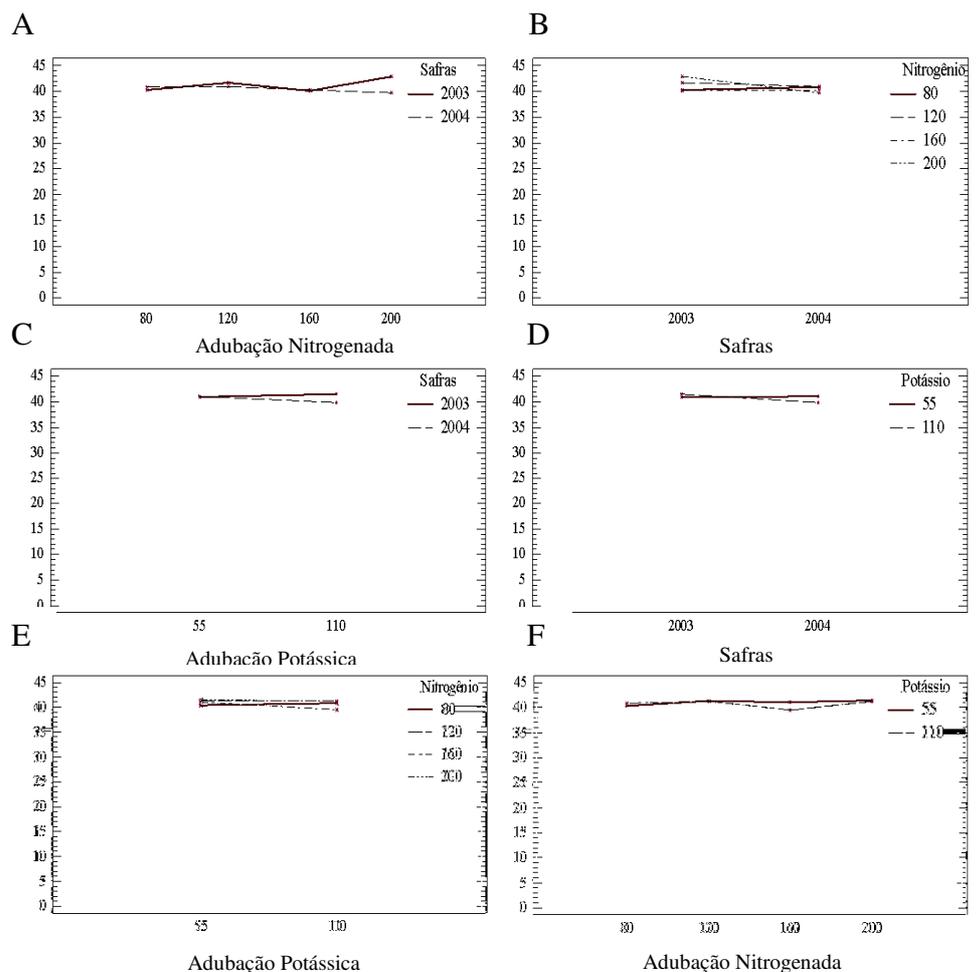


FIGURA 15 – GRÁFICA DE INTERAÇÃO DA CONCENTRAÇÃO DE NITROGÊNIO NA FOLHA (N (g Kg⁻¹)) ENTRE ADUBAÇÃO NITROGENADA, POTÁSSICA E SAFRAS NO POMAR DE PÊSSEGO ‘CHIMARRITA’, LOCALIZADO EM CAMPO LARGO, PR.

Entre as safras de 2003/04 e 2004/05 ocorreu aumento de 9,12 % no teor foliar de P, independente das adubações nitrogenada e potássica, mantendo-se no nível normal (Tabela 5, Figura 4). O resultado para P fundamentado pela significância do teste F (Tabela 3) em relação à interação entre os tratamentos de adubação nitrogenada, potássica que pode ser verificado na Figura 15. Portanto a resposta de P obtida na dose 80 kg.ha⁻¹ ano⁻¹ de N interagindo com a dose 55 kg.ha⁻¹ ano⁻¹ de K é menor quando comparada com a dose 110 kg.ha⁻¹ ano⁻¹ de K (Figura 15 A). Na dose de 160 kg.ha⁻¹ ano⁻¹ de N, o teor de P foi maior ao interagir com a dose de 55 kg.ha⁻¹ ano⁻¹ de K.

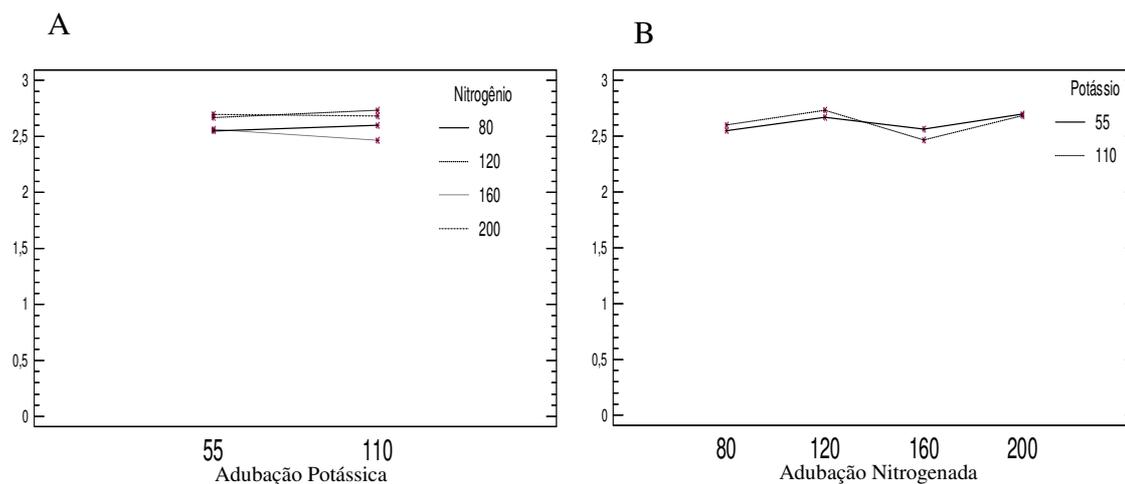


FIGURA 16 – GRÁFICOS DE INTERAÇÃO DA CONCENTRAÇÃO DE FÓSFORO NA FOLHA (P (g. Kg⁻¹)) ENTRE ADUBAÇÃO NITROGENADA E POTÁSSICA NO POMAR DE PÊSSEGO ‘CHIMARRITA’, LOCALIZADO EM CAMPO LARGO, PR.

O teor foliar K, na safra de 2003/04, aumentou 14,62% na presença de 200 kg ha⁻¹ ano⁻¹ de N. Na safra de 2004/05, o contraste significativo foi entre 80 e 160 kg.ha⁻¹ ano⁻¹ de N, este com aumento 24% no teor em relação a 80 kg ha⁻¹ ano⁻¹ de N (Figura 4, Tabela 5).

Resultados semelhantes aos dos teores foliares NPK obtidos nos pomares da Lapa e Campo Largo, foram encontrados por TEIXEIRA, SANTOS e BATAGLIA (2002), que analisaram o estado nutricional de bananeiras por meio da diagnose foliar sob três doses de adubação nitrogenada (200, 400 e 800 kg ha⁻¹ ano⁻¹) e três de potássio (300, 600 e 900 kg ha⁻¹ ano⁻¹) nas sub-parcelas e na parcela principal dois regimes hídricos: microaspersão e sequeiro, onde o tecido foliar foi amostrado de acordo com recomendações de MARTIN-PRÉVEL (1984). Outro trabalho que apresentou resultados semelhantes foi de DOLINSKI, et al., (2005) com pêssegos, avaliando os teores foliares na presença de 40, 80 e 160 kg.ha⁻¹ ano⁻¹ de adubação nitrogenada e adubação potássica (80 kg ha⁻¹ ano⁻¹) em blocos ao acaso. O estudo de MATTOS, FREIRE e MAGNANI, que analisaram teores foliares em diferentes doses de nitrogênio, obtiveram resultados semelhantes aos pomares de Campo Largo e Lapa.

TABELA 7 - ANÁLISE DOS RESULTADOS DE NITROGENIO (N g.kg⁻¹), FÓSFORO (P g.kg⁻¹) E POTÁSSIO (Kg.kg⁻¹) ENCONTRADOS NA FOLHA, EM PRESENÇA DE DIFERENTES DOSES DE ADUBAÇÃO NITROGENADA E POTÁSSICA, NO POMAR DE PÊSSEGO 'CHIMARRITA', LOCALIZADO EM CAMPO LARGO, PR.

Tratamentos Kg ha ⁻¹ ano ⁻¹	Safras	
	2003/04	2004/05
Nitrogênio		
	N (g.kg⁻¹)	
80	40,3 Aa	40,9 Aa
120	41,7 ABa	40,9 Aa
160	40,2 Aa	40,3 Aa
200	42,9 Ba	39,8 Ab
Potássio		
55	41,0 Aa	41,1 Aa
110	41,5 Aa	39,9 Ba
Nitrogênio		
	P (g.kg⁻¹)	
80	2,45 ABa	2,70 ABb
120	2,58 Ca	2,81 Bb
160	2,40 Aa	2,63 Ab
200	2,55 Bca	2,83 Bb
Potássio		
55	2,48 Aa	2,75 Ab
110	2,50 Aa	2,73 Ab
Nitrogênio		
	K (g.kg⁻¹)	
80	18,6 ABa	15,0 Aa
120	19,4 ABa	16,4 ABa
160	17,6 Aa	19,8 Ba
200	20,6 Ba	17,3 ABa
Potássio		
55	18,8 Aa	16,4 Ab
110	19,3 Aa	17,9 Aa

Teste Tukey, letras iguais minúsculas na linha e maiúsculas na coluna não diferem estatisticamente com nível de 5% significância.

Análise química realizada pela EMBRAPA - Clima temperado - Pelotas - RS, Laboratório de Nutrição Mineral de Plantas.

A análise de comparação múltipla é mais precisa que ANOVA utilizada para verificar diferenças significativas entre níveis dos fatores. Porém, se faz necessário a utilização de gráficos quando as interações entre os fatores são significativas.

A análise dos teores foliares é composta por mais de um elemento, ou seja, variável resposta. A ANOVA não é capaz de fornecer resultados que leva em consideração nas análises a interação entre as variáveis respostas. Portanto um método mais eficiente que ANOVA é Modelos Lineares Generalizados.

3.8 ANÁLISE DE MODELOS LINEARES GENERALIZADOS

As tabelas 6 e 7 apresentam as análises do Modelo Linear Generalizado (GLM) para os teores foliares NPK, ajustado pela Distribuição Probabilística Normal, juntamente com as estimativas dos parâmetros que ajudam a explicar as variáveis respostas NPK encontrados nas folhas nos pomares da Lapa e Campo Largo, conforme a significância do teste.

Lapa. O teor foliar de N foi influenciado positivamente pela concentração $160 \text{ kg.ha}^{-1} \text{ ano}^{-1}$ de N, em segundo por $200 \text{ kg.ha}^{-1} \text{ ano}^{-1}$ de N, entretanto, o terceiro bloco e a safra de 2004/05 influenciaram negativamente na variável resposta de N. Dentre os macros e micronutrientes, o que mais influenciou o teor foliar de N foi o P, sendo este o principal fator que explica a quantidade de N encontrado na folha (Tabela 6).

Os parâmetros que ajudam a explicar a quantidade de K encontrado na folha são os blocos 2, 3 e a safra 2005/06, este ultimo é o fator que mais explica o teor foliar de K. Dentre os macros e micronutrientes os que mais influenciam foram Ca positivamente e Mg negativamente (Tabela 6).

A quantidade de P encontrado nas folhas esta fracamente dependente da safra 2005, os demais fatores, macro e micronutrientes não ajudam, acentuadamente, a explicar o teor foliar de P (Tabela 6).

TABELA 8 - MODELO LINEAR GENERALIZADO AJUSTADO PARA AS VARIÁVEIS, ENCONTRADA NAS FOLHAS, NITROGENIO (N g.kg⁻¹), FÓSFORO (P g.kg⁻¹) E POTÁSSIO (K g.kg⁻¹) NA PRESENÇA DOS TEORES FOLIARES DE MACRO E MICRONUTRIENTES E ADUBAÇÃO NITROGENADA E POTÁSSICA, NOS POMARES DE PÊSSEGO, LOCALIZADO NA LAPA-PR.

β Parâmetro	Estimativa	Erro padrão	valor t	P-valor
Variável N				
β_0 Intercepto	-5.19	4.977	-1.043	0.299
β_1 P (g.kg ⁻¹)	13.83	2.240	6.173	2.35e ⁻⁸
β_2 Mn (mg.kg ⁻¹)	0.048	0.003	12.59	< 2e ⁻¹⁶
β_3 160 kg.ha ⁻¹ Ano ⁻¹ de N	2.323	0.506	4.586	1.58e ⁻⁵
β_4 200 kg.ha ⁻¹ Ano ⁻¹ de N	1.368	0.491	2.785	0.006
β_5 3º Bloco	-1.59	0.399	-3.985	0.001
β_6 Safra de 2004	-1.813	0.408	-4.442	2.73e ⁻⁵
Variável K				
β_0 Intercepto	23.00	1.594	14.421	< 2e ⁻¹⁶
β_1 Ca (g.kg ⁻¹)	0.449	0.099	4.519	2.04e ⁻⁵
β_2 Mg (g.kg ⁻¹)	-1.218	0.339	-3.593	0.0005
β_3 Fe (g.kg ⁻¹)	-0.037	0.006	-5.552	3.31e ⁻⁷
β_4 2º Bloco	0.838	0.281	2.978	0.003
β_5 3º Bloco	0.713	0.291	2.449	0.016
β_6 Safra de 2005	1.743	0.372	4.679	1.11e ⁻⁵
Variável P				
β_0 Intercepto	1.770	0.076	23.178	< 2e ⁻¹⁶
β_1 N (g.kg ⁻¹)	0.021	0.002	7.338	1.19e ⁻¹⁰
β_2 Mn (g.kg ⁻¹)	-0.001	0.000	-4.618	1.36e ⁻⁰⁵
β_3 3º Bloco	0.046	0.016	2.917	0.0045
β_4 Safra de 2005	-0.09	0.023	-3.989	0.0001

* p-valor < 0,05 é significativo

Campo Largo. O teor foliar de N foi influenciado positivamente pela interação entre safra de 2004/05 e a concentração 200 kg.ha⁻¹ ano⁻¹ de N, este com uma maior influência do que a safra. Dentre os macros e micronutrientes, o que mais influenciou, negativamente, o teor foliar de N foi Ca (Tabela 7).

Os parâmetros que ajudaram a explicar a quantidade de K encontrado nas folhas foram os blocos I, II e a interação entre safra 2005/06 e 160 kg.ha⁻¹ ano⁻¹ de N, este ultimo é o fator que mais explica o teor foliar de K dentre a interação. Dentre os macros e micronutrientes os que mais influenciam foram Ca positivamente e Mg negativamente. O fator que melhor ajuda explicar a quantidade de K é a interação entre 160 kg.ha⁻¹ ano⁻¹ de N e Safra 2004/05 (Tabela 7).

A quantidade de P encontrado na folha esta fracamente dependente da safra 2003, os demais fatores, macro e micronutrientes não ajudam, acentuadamente, a explicar o teor foliar de P (Tabela 7).

TABELA 9 - MODELO LINEAR GENERALIZADO AJUSTADO PARA AS VARIÁVEIS, ENCONTRADA NAS FOLHAS, NITROGÊNIO (N g.kg⁻¹), FÓSFORO (P g.kg⁻¹) E POTÁSSIO (K g.kg⁻¹) NA PRESENÇA DOS TEORES FOLIARES DE MACRO E MICRO NUTRIENTES E ADUBAÇÃO NITROGENADA E POTÁSSICA, NOS POMARES DE PÊSSEGO, LOCALIZADO EM CAMPO LARGO-PR.

β Parâmetro	Estimativa	Erro padrão	valor t	P-valor
Variável N				
β ₀ Intercepto	47.984	2.350	20.415	< 2e ⁻¹⁶
β ₁ Ca (g.kg ⁻¹)	-0.352	0.112	-3.131	0.003
β ₂ 200 kg.ha ⁻¹ Ano ⁻¹ de N	1.7195	0.673	2.553	0.014
β ₃ Safra de 2004	-0.443	0.486	-0.910	0.367
β ₄ 200 de N * Safra 2004	-2.625	0.943	-2.782	0.007
Variável K				
β ₀ Intercepto	19.548	4.336	4.508	5.6e ⁻⁵
β ₁ Ca (g.kg ⁻¹)	0.5943	0.209	2.838	0.007
β ₂ Mg (g.kg ⁻¹)	-2.543	0.932	-2.729	0.009
β ₃ 1° Bloco	-2.819	0.817	-3.449	0.001
β ₄ 2° Bloco	-2.088	0.720	-2.898	0.006
β ₅ 160 kg.ha ⁻¹ Ano ⁻¹ de N	-2.229	0.960	-2.321	0.025
β ₆ Safra de 2005	0.9730	1.644	0.592	0.557
β ₇ 160 de N * Safra 2004	4.7834	1.434	3.336	0.001
Variável P				
β ₀ Intercepto	2.7458	0.02387	115.021	< 2e ⁻¹⁶
β ₁ Safra de 2003	-0.250	0.03376	-7.405	2.26e ⁻⁹

* p-valor < 0,05 é significativo

Com base nos resultados apresentados na ANOVA de ambos os pomares, somente no pomar de Campo Largo o fator bloco foi significativo. Entretanto, por meio das análises com GLM, no pomar da Lapa o fator bloco ajuda a explicar o teor foliar NPK, e no pomar de Campo Largo o teor foliar K, isso evidencia a necessidade de se utilizar blocagem no delineamento experimental. Estas dependências também podem ser observadas nas análises de correlações realizadas anteriormente. Além disso, o GLM possibilita facilmente a verificação dos macro e micronutrientes (variáveis explicativas) que influenciaram em NPK (variáveis respostas).

Em ambos os pomares a adubação potássica não ajudou a explicar a quantidade de NPK, corroborando com as análises de correlações de Spearman realizadas anteriormente (Tabela 6, 7).

Na identificação das interações entre os fatores, as análises GLM e ANOVA apresentaram em seus resultados uma dependência entre as safras e adubação nitrogenada no pomar de Campo Largo.

No trabalho, realizado por DOLINSKI, (2007), com adubação nitrogenada e potássica na cultura de ameixeira “*Reubennel*” na região de Araucária, adubação 55 a 110 kg.ha⁻¹ ano⁻¹ de K não influenciaram no teor de K encontrado na folha. Verificou que os resultados obtidos neste trabalho, esta fundamentado no método estatístico ANOVA com delineamento parcelas sub-subdividida. Este trabalho pode ser otimizado utilizando métodos de GLM com distribuição Gauss, considerando os demais teores foliares disponíveis. Para utilizar GLM nas demais variáveis disponíveis no estudo de DOLINSKI, (2007) deve ser verificado qual é a melhor distribuição probabilística na modelagem.

4 CONCLUSÕES

- a. Desenhos esquemáticos se mostraram simples e extremamente úteis para identificação de valores discrepantes em macro e micronutrientes.
- b. A análise GLM e de Correlação de Spearman foram eficientes para identificar a necessidade de utilizar blocagem.
- c. ANOVA não foi eficiente para identificar a necessidade de blocagem.
- d. Análises GLM e ANOVA apresentaram as mesmas conclusões referentes à dependência entre safras e as doses de nitrogênio.
- e. Análise de GLM é mais eficiente que ANOVA para analisar mais de uma variável resposta simultaneamente.
- f. Os macros e micronutrientes estão relacionados entre as safras.
- g. Adubação nitrogenada 160 e 200 kg.ha⁻¹ ano⁻¹ influenciaram na quantidade de N encontrado na folha.
- h. A adubação potássica não influenciou nos teores foliares NPK.

REFERÊNCIAS

ABREU, M. A.; VETTER, D. A análise de relações entre conjuntos de variáveis na matriz geográfica: correlação canônica. In: FAISSOL, S. **Tendências atuais na geografia urbano/regional: teorização e quantificação**. Rio de Janeiro: IBGE, 1978. p. 133-144.

ALVES, H.; CUNHA, M. L. **Software Estatístico – Uma introdução a alguns aplicativos, numa abordagem inicial dos dados**. Dossiê Didático. Disponível em: <<http://alea-esp.ine.pt>>. Acessado em 22 de Dezembro de 2007.

ASSMANN, S. T. **Influência da aplicação do calcário Irati (São Mateus do sul – PR) nos teores de S, cu, Fe, Mn e Zn no Solo e na Planta de Milho (*Zea may L.*)**. Curitiba, 1995, Dissertação (Mestrado em Ciências do Solo) - Universidade Federal do Paraná – UFPR, 83p

BARBIN, D. **Componentes de variância: teoria e aplicações**. Piracicaba: FEALQ, 2004. 120p.

BARCIK, C. **Processos Autoalelopáticos na cultura de Alfafa (*Medicago sativa L. CV Crioula*) em solos de diferentes texturas**. Curitiba, 1999, Dissertação (Mestrado em Ciências do Solo) - Universidade Federal do Paraná – UFPR, 109p.

CHRISTENSEN, R.; BEDRICK, J, E. Testing the Independence Assumption in Linear Models. **Journal of the American Statistical Association**, v. 92, n. 439, p. 1006-1016, 1997.

BATAGLIA, O. C; SANTOS, R. W. Efeitos da Adubação NPK na Fertilidade do Solo, Nutrição e Crescimento da Seringueira. **Revista Brasileira de Ciências do Solo**, Campinas, v. 23, n. 4, p. 881-890, 1999.

BENJAMINI, Y. Opening the Box of a Boxplot. **The American Statistician**, Alexandria, v. 42, n. 4, p. 257-262, 1988.

BEST, D. J.; RAYNER, J. C. W. **Smooth extensions of Pearsons's product moment correlation and Spearman's rho**. Estados Unidos: *Statistics & Probability Letters*, 1996. pag. 171 – 177.

BUSSAB, W. O.; MORETTIN, P. A. **Estatística Básica**. São Paulo: Atual, 1987. 436 p.

CAMPOS, G. M.; BARBIN, L. E.; SPANÓ, E. C. J.; SILVA, S. R.; PÉCORA, D. J. **Estatística Prática para Docentes e Pós-Graduandos: tipos de variáveis**. Disponível em: <http://www.forp.usp.br/restauradora/gmc/gmc_livro/gmc_livro_cap02.html>. Acesso em: 23 de setembro de 2007.

CAVADIAS, G.; PILON, P.; YUE, S. Power of the Mann-Kendall and Spearman's rho tests for monotonic trends in hidrological series. **Journal of Hydrology**, New Zealand, v. 259, n. 1, p. 254 – 271, 2002.

COMISSÃO DE FERTILIDADE DO SOLO RS/SC. **Recomendações de adubação e de calagem para os estados do Rio Grande do Sul e de Santa Catarina**. 3 ed. Passo Fundo, Sociedade Brasileira de Ciência do Solo, 1995. 224p.

COSTA, N. P. O. **Estatística**. São Paulo: Edgard Blucher, 1977.

DEMETRIO, C. G. B. **Modelos lineares Generalizados em experimentação agrônômica**. Piracicaba: ESALQ, 2001. 113p.

DOLINSKI, M. A. **Adubação nitrogenada e potássica na cultura de ameixeira 'Reubennel' na região de Araucária – PR**. Curitiba, 2007, Dissertação (Mestrado) - Universidade Federal do Paraná – UFPR.

DOLINSKI, M. A.; SERRAT, B. M.; MOTTA, A. C. V.; CUQUEL, F. L.; SILVANA REGINA DE SOUZA, S. R.; MAY-DE MIO, L. L.; MONTEIRO, L. B. Produção, teor foliar e qualidade de frutos do pessegueiro 'chimarrita' em função da adubação nitrogenada, na região da Lapa - PR. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 27, n. 2, p. 295-299, 2005.

ESTELBERGER, W; REIBNEGGER, G. The rank correlation coefficient: na additional AID in the interpretation of laboratory data. **Journal Clinica Chimica Acta**, Amsterdam, v. 239, n. 2, p. 203-207.

FACHINELLO, J. C.; COUTINHO, E. F.; MARODIN, G. A.; BOTTON, M.; MAY DE MIO, L. L. **Normas técnicas e documentos de acompanhamento da produção integrada de pêssego**. Pelotas: Signus, 2003. 92p.

FAYYAD, U. M. **Advances in Knowledge Discovery and Data Mining**. Menlo Park: AAAI Press/MIT Press, 1996. 611p.

GIOLO, S. R. **Análise de Regressão e Correlação**. Curitiba: Universidade Federal do Paraná. 2005.

GONÇALVES, E. B.; CASTRO, I. M. Desenho esquemático para programas de qualidade em laboratórios de análises. **Ciência e Tecnologia de Alimentos**, Campinas, v. 18, n. 1, p. 121 – 126, 1998.

HAIR, J. F.; ANDERSON, R. E.; TATHAM, R. L.; BLACK, W. C. **Multivariate Data Analysis**. New Jersey: Prentice Hall, 1995.

HOAGLIN, D.C.; MOSTELLER, F. & TUKEY, J.W. **Understanding robust and exploratory analysis**. New York: Wiley, 1983. 447 p.

LAHAV, E. Banana nutrition. In: GOWEN, S. (Ed.) **Bananas and plantains**. London: Chapman & Hall, 1995. p. 258-316.

MATTAR, F. N. **Pesquisa de marketing: metodologia, planejamento, execução e análise**. 2. ed. São Paulo: Atlas, 1994. 275 p.

MATTOS, M. L. T.; FREIRE, C. J. S.; MAGNANI, M. Crescimento e teores foliares de N, P, K, Ca e Mg em pessegueiro cv. Diamante com diferentes níveis de N aplicado ao solo. **Pesquisa Agropecuária Brasileira**, Brasília, v. 26, n. 9, p.1315-1321, 1991 a.

MATTOS, M. L. T.; FREIRE, C. J. S.; MAGNANI, M. Produção do pessegueiro cv. Diamante, sob diferentes doses de nitrogênio aplicado ao solo. **Pesquisa Agropecuária Brasileira**, Brasília, v. 26, n. 1, p.113-117, 1991b.

MARTIN-PRÉVEL, P. B.; GAGNARD, J., GAUTIER, P. **L'analyse végétale dans Le contrôle de l'alimentation des plantes tempérées et tropicales**. Paris: Tec&Doc, 1984. p.715-51.

MARTIN-PRÉVEL, P. Past, present and future of tropical fruit nutrition with special reference to banana nutrition. **Acta Horticulturae**, Leuven, v. 275, p. 523-34, 1990.

MCCULLAGH, P.; NELDER, J. A. **Generalized linear models**. 2 ed. London: Chapman & Hall, 1989. 511 p.

MONTGOMERY, D. C. **Design and analysis of experiments**. 3 ed. New York: John Wiley & Sons, 2001. 856 p.

NETO, J. M. M. **Estatística multivariada, Uma visão didática-metodológica**. Disponível em <http://criticanarede.com/cien_estatistica.html>. Acesso em 23 de fevereiro de 2007.

PEARSON, R. K. Exploring process data. **Journal of Process Control**, Maryland, v. 11, n. 2, p. 79-194, 2001.

PIMENTEL-GOMES, F.; VALSECHI, O.; ABREU, C. P; OLIVEIRA, E. R. Amostragem da cana-de-açúcar para determinações tecnológicas. **Anais da Escola Superior de Agricultura Luíz de Queiroz**, Piracicaba, v.20, p.89-114, 1963.

PONTES, F. C. A. **Análise de variância multivariada com a utilização de métodos não-paramétricos e componentes principais baseados em matrizes de postos**. Piracicaba, 2005. Tese (Doutorado) – ESALQ, Universidade de São Paulo.

R DEVELOPMENT CORE TEAM. R. **Writing R extensions**. Disponível: <<http://cran.r-project.org/doc/manuals/R-exts.pdf>> Acesso em: 01 de fevereiro de 2006.

R DEVELOPMENT CORE TEAM. R. **A language and environment for statistical Computing**. R Foundation for Statistical Computing, Vienna. Disponível em: <<http://www.R-project.org>> Acesso em: 01 de fevereiro de 2006.

SCHUHMACHER, F. E. **Avaliação das características agronômicas e nutricionais de dois cultivares de milho (Zea mays L.) sob diferentes níveis de calagem**. Curitiba, 1997, Dissertação (Mestrado) - Universidade Federal do Paraná – UFPR.

SCHWERTMAN, C. N.; SILVA, R. **Identifying outliers with sequential fences**. *Computational Statistics & Data Analysis* v. 51, n. 8, p. 3800 – 3810, 2007.

SIEGEL, S.

SILVA, F. C. **Manual de Análises Químicas de Solo, Planta e Fertilizante**. Brasília: Embrapa Solos, 1999. p. 175-222.

SIT, V. **Catalogo f curves for curves fitting: Handbook 4**. Victoria: Ministry of Forest, Forest Science Research Branch, Biometrics information, 1994.

SOUZA, R. S. **Adubação Nitrogenada no Desenvolvimento das Doenças do Pessegueiro, sob Sistema de Produção Integrada de Fruteiras, na Lapa, Pr**. Curitiba, 2005, Dissertação (Mestrado) - Universidade Federal do Paraná – UFPR.

SPSS Inc. **SPSS statistical software**. Version 12.0, 2003. Illinois, 2003.

TEIXEIRA L. A. J., SANTOS W. R. D., BATAGLIA O. C. Diagnose nutricional para nitrogênio e potássio em bananeira por meio do sistema integrado de diagnose e recomendação (DRIS) e de níveis críticos. **Revista Brasileira de Fruticultura**, Jaboticabal, v. 24, n. 2, p. 530-535, 2002.

TOLEDO, G. L.; OVALLE, I. I. **Estatística Básica**. São Paulo: Editora: Atlas, 1995. p. 430-431.

TREVISAN, R.; TREPTOW, O. R.; GONÇALVES, D. E.; ANTUNES, C. E. L.; HERTER, G. F. Influência do manejo fitotécnico na qualidade sensorial de pêssego (*Prunus persica* (L.) Batsch) cv. Maciel. **Revista Brasileira Agrociência**, Pelotas, v. 12, n. 4, p. 409-414, 2006.

TRUGILHO, P. F.; LIMA, T. J.; MORI, A. F. Correlação canônica das características químicas e físicas da madeira de clones de *Eucalyptus grandis* e *Eucalyptus saligna*. *Cerne*, Lavras, v. 9, n. 1, p. 066 - 080, 2003.

TUKEY, J. W. **Exploratory data analysis**. Reading: Addison-Wesley, 1977. 588 p.

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)