
MÉTODOS DE SELEÇÃO DE EQUAÇÕES DE REGRESSÃO LINEAR MÚLTIPLA.

CARLOS HENRIQUE MATTIOLLI

Eng. Agr. - EMBRAPA

Orientador: Dr. ROBERTO SIMIONATO MORAES

Dissertação apresentada à Escola Superior de Agricultura "Luiz de Queiroz", da Universidade de São Paulo, para obtenção do título de Mestre em Agronomia. Área de Concentração: Estatística e Experimentação Agronômica.

PIRACICABA
Estado de São Paulo - Brasil
Junho, 1983

A Maria Aparecida,

minha esposa,

A meus pais e aos meus sogros,

DEDICO.

AGRADECIMENTOS

- Ao Dr. Roberto Simionato Moraes, Professor Assistente Doutor do Departamento de Matemática e Estatística da ESALQ, pela valiosa orientação e revisão dos originais.
- Ao Dr. Humberto de Campos, Professor Titular do Departamento de Matemática e Estatística da ESALQ, e Coordenador do Curso de Pós-Graduação em Estatística e Experimentação Agronômica, pelo apoio concedido no decorrer do Curso.
- Ao Dr. José Claret Matioli, Gerente de Pesquisas da Divisão Agroquímica da Asfaltos Chevron S.A., pela cessão dos dados experimentais e pelas valiosas sugestões.
- Ao Dr. Ricardo Borher Sgrillo, Pesquisador do Centro de Energia Nuclear na Agricultura, pelas sugestões e leitura do texto.
- Aos Professores do Departamento de Matemática e Estatística da ESALQ, pelos ensinamentos.
- Aos colegas Ariovaldo Miguel Carvalho e Francisco Carlos Pelegrino, pelo apoio e amizade.
- Ao colega Marcos Amaral Manfrinato, pelo desenho do fluxograma.
- Ao colega Valter João Diehl, pelas sugestões e apoio.
- A todos os pesquisadores e funcionários do CENA, pelo apoio e serviços prestados.

- A todos os funcionários do Departamento de Matemática e Estatística da ESALQ, pelos serviços prestados.
- À Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA), pela oportunidade de frequentar este Curso.
- Ao Centro de Energia Nuclear na Agricultura, pelas excelentes condições de trabalho oferecidas.
- A todos que de alguma forma colaboraram para a realização deste Curso e deste trabalho.

Í N D I C E

	Pág.
RESUMO	xi
SUMMARY	xiv
1. INTRODUÇÃO	1
2. REVISÃO DE LITERATURA	5
2.1 - Formulação do Problema	5
2.2 - Avaliação de todas as Possíveis Regressões	7
2.2.1 - Generalidades	7
2.2.2 - Algoritmos	8
2.2.3 - Métodos de fracionamento	9
2.3 - A Multicolinearidade	10
2.3.1 - Generalidades	10
2.3.2 - Sintomas	11
2.3.3 - Remédios	13
2.3.4 - Restrição à utilização dos métodos STEPWISE	13
2.4 - Critérios para a Seleção de Variáveis	14
2.4.1 - Generalidades	14
2.4.2 - A Estatística C_p de MALLOWS (1964)	15
2.4.2.1 - Definição	15
2.4.2.2 - Derivação	15
2.4.2.3 - O método gráfico	17

	Pág.
2.5 - Análise de Resíduos	19
2.5.1 - Generalidades	19
2.5.2 - Métodos gráficos	20
2.5.3 - Estatística para exame dos resíduos	21
3. MATERIAL E MÉTODOS	24
3.1 - Material	24
3.2 - Métodos	30
3.2.1 - Introdução	30
3.2.2 - Ajustamento das equações completas e avaliação da multicolinearidade	31
3.2.3 - O método para avaliação de todas as regressões possíveis	32
3.2.3.1 - O algoritmo proposto	32
3.2.3.2 - O cálculo dos C_p 's	37
3.2.4 - Análise gráfica dos resíduos	37
3.2.5 - O teste dos chorrilhos para autocorrelação dos resíduos	38
3.2.6 - O teste de Lilliefors para normalidade dos resíduos	41
3.2.7 - Determinação de modelos pelo método dos componentes principais	43
4. RESULTADOS E DISCUSSÃO	45
5. CONCLUSÕES	84
6. BIBLIOGRAFIA	87

LISTA DE TABELAS

TABELA		Pág.
1	Dados de avaliações de danos em grãos de milho C-111	27
2	Dados de avaliações de danos em grãos de milho FLINT	28
3	Dados de avaliações de danos em grãos de milho PIRANÃO	29
4	Estatísticas obtidas para a equação completa referente ao cultivar C-111	47
5	Estatísticas obtidas para a equação completa referente ao cultivar FLINT	48
6	Estatísticas obtidas para a equação completa referente ao cultivar PIRANÃO	49
7	Autovalores obtidos para as matrizes de correlações	50
8	Todas as possíveis regressões fornecidas pelo programa MALLOW'S para o cultivar C-111	52
9	Todas as possíveis regressões fornecidas pelo programa MALLOW'S para o cultivar FLINT	53
10	Todas as possíveis regressões fornecidas pelo programa MALLOW'S para o cultivar PIRANÃO	54
11	Vinte primeiras equações em ordem crescente de C_p 's	55

TABELA	Pág.
12 Resultados do programa STEP.COMP, para análise de regressão da equação ACE, selecionada para o cultivar C-111	58
13 Resultados do programa STEP.COMP, para análise de regressão da equação ADH, selecionada para o cultivar FLINT	59
14 Resultados do programa STEP.COMP, para análise de regressão da equação ABD, selecionada para o cultivar PIRANÃO	60
15 Distribuição nula dos resíduos para os três cultivares de milho, geradas pelo programa STEP.COMP	64
16 Resultados do teste de Lilliefors para normalidade dos resíduos do cultivar C-111, fornecidos pelo programa STEP.COMP	66
17 Resultados do teste de Lilliefors para normalidade dos resíduos do cultivar FLINT, fornecidos pelo programa STEP.COMP	67
18 Resultados do teste de Lilliefors para normalidade dos resíduos do cultivar PIRANÃO, fornecidos pelo programa STEP.COMP	68
19 Componentes principais para os três cultivares de milho	71

TABELA		Pág.
20	Autovalores, autovetores e matriz ortogonal de escores do cultivar C-111, fornecidos pelo programa AUTO-JACOBI	72
21	Autovalores, autovetores e matriz ortogonal de escores do cultivar FLINT, fornecidos pelo programa AUTO-JACOBI	73
22	Autovalores, autovetores e matriz ortogonal de escores do cultivar PIRANÃO, fornecidos pelo programa AUTO-JACOBI	74
23	Resultados das análises de componentes principais dos três cultivares de milho	75
24	Resultados da análise de regressão dos componentes principais selecionados para o cultivar C-111	76
25	Resultados da análise de regressão dos componentes principais selecionados para o cultivar FLINT	77
26	Resultados da análise de regressão dos componentes principais selecionados para o cultivar PIRANÃO	78
27	Distribuições nulas dos resíduos, provenientes dos modelos de componentes principais, para os três cultivares de milho	80

TABELA		Pág.
28	Resultados do teste de Lilliefors para normalidade dos resíduos do cultivar C-111, provenientes do modelo de componentes principais	81
29	Resultados do teste de Lilliefors para normalidade dos resíduos do cultivar FLINT, provenientes do modelo de componentes principais	82
30	Resultados do teste de Lilliefors para normalidade dos resíduos do cultivar PIRANÃO, provenientes do modelo de componentes principais	83

MÉTODOS DE SELEÇÃO DE EQUAÇÕES DE REGRESSÃO
LINEAR MULTIPLA

Candidato: Carlos Henrique Mattioli

Orientador: Prof. Dr. Roberto Simionato Moraes

R E S U M O

Utilizando dados relativos a danos físicos e químicos causados por *Sitophilus oryzae* em grãos de milho armazenados, foram determinadas equações para estimação de suas perdas de peso, através do cálculo de todas as possíveis regressões e dos componentes principais, objetivando demonstrar a importância da realização de análises mais refinadas na obtenção de equações mais confiáveis.

Inicialmente foi verificada a existência de multicolinearidade para se determinar o método e o critério de seleção de variáveis mais convenientes. Esta escolha recaiu sobre o método de avaliação de todas as possíveis regressões e a estatística C_p própria ao mesmo, pela sua aplicabilidade em dados colineares, e por prover maior volume de informações.

Ao contrário de se buscar somente equações com elevados coeficientes de determinação, os ajustamentos foram realizados levando-se em consideração a distribuição dos resíduos através de estudos gráficos e analíticos dos mesmos, verificando-se possíveis violações nas hipóteses básicas de independência, homocedasticidade e normalidade dos erros.

Foram desenvolvidos os algoritmos e programas em linguagens BASIC para total processamento dos cálculos em microcomputador.

Tendo sido constatada acentuada multicolinearidade, foram feitos ajustamentos de equações pelo método dos componentes principais, indicado para estas circunstâncias, extraíndo-se e estudando seus resíduos.

Os resultados obtidos nos fizeram concluir que:

- 1) Os métodos propostos, embora levando a equações com elevada explicabilidade, não foram adequados por apresentarem estimativas tendenciosas para as perdas de peso.
- 2) As seleções de todas as possíveis regressões, embora não adequadas, forneceram informações relevantes sobre a importância das variáveis estudadas.
- 3) As análises de resíduos foram imprescindíveis nas considerações sobre a inadequação das equações, demonstrando acentuados comportamentos não lineares dos mesmos.

- 4) A utilização dos programas desenvolvidos em linguagem BASIC para microcomputadores conduziu a bons resultados quanto à capacidade, precisão e tempo de processamento.

SELECTION METHODS OF MULTIPLE LINEAR
REGRESSION EQUATIONS

Candidate: Carlos Henrique Mattioli
Adviser: Prof. Dr. Roberto Simionato Moraes

S U M M A R Y

Using data related to physical and chemical damage caused by *Sitophilus oryzae* in stored maize grains, equations have been determined to estimate loss in weight, by calculating all possible regressions and main components, with the objective of proving the importance of using more sophisticated analysis to obtain more reliable equations.

The existence of multicollinearity was initially checked to determine the method and selection criterium of more convenient variables. The method chosen was the evaluation of all possible regressions and the corresponding appropriate C_p statistics, due to its applicability to colinearity and as a greater source of information.

Instead of trying to find equations with high determination coefficients, adjustments were made taking into account the distribution of residues through graphic and analytical studies, which showed possible disturbance in the basic assumption of independence, homocedasticity and normality.

Algorithms and programmes were developed in BASIC language to process microcomputer calculations. Since high multicollinearity was noticed, the equations were adjusted using the main components method, recommended under these circumstances, the residues being extracted and studied.

In view of the results obtained, it was concluded:

1. the proposed methods, although leading to highly explainable equations were not adequate, because of the presence of bias in the estimates of weight loss;
2. the selection of all possible regressions, although not adequate, supplied relevant information on the importance of the variables studied;
3. the analysis of the residues, indicating marked non-linear behaviour, was essential in the determination of the inadequacy of the equations;
4. The utilization of programmes for microcomputers, developed in BASIC language, has led to good results as far as capacity, accuracy and processing time are concerned.

1. INTRODUÇÃO

A análise de regressão, uma das ferramentas estatísticas mais amplamente utilizadas, com aplicações nos mais diversos ramos da Ciência, consiste em um conjunto de técnicas para estudo de relacionamentos entre variáveis, de grande interesse teórico pela elegância de sua teoria, e também prático pela metodologia relativamente simples.

O relacionamento entre variáveis pode ser expresso na forma de uma equação a qual associa uma variável dependente Y , a uma ou mais variáveis independentes X_1, X_2, \dots, X_p , tendo, na forma linear a seguinte expressão:

$$Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_pX_p \quad ,$$

onde $b_0, b_1, b_2, \dots, b_p$ são chamados de coeficientes de regressão

e estimados através dos dados . A equação representada na forma $Y = b_0 + b_1X$ denomina-se equação de regressão linear simples. A equação representada na forma $Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_pX_p$ denomina-se equação de regressão linear múltipla.

As razões pelas quais procuramos ajustar uma equação de regressão são variadas, mas, de uma maneira geral, as mais comuns são para estabelecer relacionamentos funcionais entre variáveis, e também fazer previsões.

Se dispomos de um conjunto de K variáveis independentes, das quais desejamos extrair uma equação de regressão linear, contudo sem sabermos ao certo quais destas variáveis são realmente importantes e necessárias na sua construção, podemos lançar mão de alguns dos métodos de seleção de variáveis.

Fundamentalmente, existem dois tipos destas técnicas. A primeira, mais trabalhosa porém mais conclusiva, consiste no cálculo de todas as regressões possíveis. O outro tipo, importante e de filosofia diferente, é chamado de regressão "stepwise", ou regressão passo a passo. Ao contrário de calcular todas as regressões possíveis, esta técnica procura traçar um caminho ótimo entre elas, verificando as mudanças causadas na explicabilidade da equação, pela inclusão e exclusão de variáveis na mesma.

São três os algoritmos básicos para regressão "stepwise":

1. "FORWARD SELECTION" (FS). As variáveis independentes são incluídas uma a uma na equação. A cada etapa, a variável X_i que promover um decréscimo mais significativo na soma de quadrados residual, e cujo valor calculado de F , dado por $|\hat{\beta}_i/s(\hat{\beta}_i)|^2$, exceder a um valor F pré-estabelecido, será incluída na equação.

2. "BACKWARD ELIMINATION" (BE). Iniciando com a equação completa, esta versão vai eliminando uma a uma as variáveis menos importantes, utilizando um critério semelhante ao utilizado na versão FS.

3. "STEPWISE" (SW). Basicamente é a versão FS, implementada com a possibilidade de excluir variáveis que deixam de ser importantes à medida que novas variáveis são incluídas na equação.

Um aspecto de grande importância, mas que nem sempre recebe a devida atenção, é a existência de acentuadas correlações entre as variáveis independentes, notadamente conhecidas como multicolinearidade. Nestas circunstâncias, a utilização dos métodos "stepwise" pode culminar em resultados confusos. Nestes métodos é implícita a suposição de que existe uma única "melhor" equação e que eles irão determiná-la. Isto nem sempre ocorre, pois, em diversas situações podem existir equações inclusive melhores, constituídas por outras diferentes combinações de variáveis, e que são negligenciadas por estes métodos.

Na presença de multicolinearidade, a associação de variáveis em diferentes ordens pode levar a diferentes equações fi

nais, tornando arriscada a seleção de variáveis pela estatística F , usualmente utilizada como critério de seleção nos métodos "stepwise".

Utilizando três conjuntos de variáveis, as quais são avaliações de danos físicos e químicos causados a grãos de milho armazenados, pela infestação de *Sitophilus oryzae*, o presente trabalho teve como objetivos:

1. Demonstrar a utilização do método de seleção de variáveis independentes pela avaliação de todas possíveis regressões, e a aplicação da estatística C_p definida na página 15 como critério de seleção em lugar da usual estatística F .

2. Ressaltar a importância de uma análise dos resíduos na escolha da equação final.

3. Elaborar programas para utilização em microcomputadores que nos permitam o processamento mais refinado obtendo consequente maior precisão dos resultados.

2. REVISÃO DE LITERATURA

2.1 - Formulação do Problema

Para WEISBERG (1980), a principal e mais importante ferramenta na seleção de variáveis em análises de regressão é um substancial conhecimento da área e de cada variável em estudo, incluindo o sinal e a magnitude dos parâmetros.

Já AFIFI e AZEN (1979) comentam que nem sempre é possível ao pesquisador ter informações sobre a ordem de importância, quantas e quais variáveis deverão fazer parte da equação.

DANIEL e WOOD (1980) relatam que, em suas experiências, frequentemente são observadas diferentes equações igualmente representativas dos dados, recomendando que outros aspectos devem ser levados em consideração na determinação da "melhor" equação.

Os seguintes aspectos são considerados importantes por CHATTERJEE e PRICE (1977):

1) A finalidade a qual se destina a equação.

Uma equação cujo propósito é a descrição de fenômenos ou processos deverá ser construída de maneira tal que um menor número de variáveis independentes consiga explicar a parte mais substancial das variações na variável dependente ou resposta. Em situações onde a predição é o objetivo, as variáveis deverão ser selecionadas com vistas a minimizar o quadrado médio residual, o que nos leva a desejar equações com um maior número de variáveis independentes.

2) Os custos envolvidos na obtenção de um maior número de variáveis.

3) A facilidade de obtenção das variáveis.

Se duas variáveis são similarmente relacionadas com a variável resposta, podemos optar pela utilização daquela de mais fácil obtenção, mesmo que o seu grau de correlação seja inferior, pois nem sempre a variável mais correlacionada torna-se mais importante na equação final.

DRAPER e SMITH (1966), DANIEL e WOOD (1980) e BASKERVILLE e TOOGOOD (1982) apresentam, com diferentes enfoques, organogramas a serem seguidos na construção de modelos de regressão linear múltipla.

2.2 - Avaliação de todas Possíveis Regressões

2.2.1 - Generalidades

Dado um modelo $Y = b_0 + b_1X_1 + \dots + b_kX_k + \text{erro}$ são possíveis os ajustamentos de 2^k equações de regressão, sendo $Y = \hat{b}_0$ ($\hat{b}_0 = \bar{Y}$) a mais simples delas. Além desta, temos:

- $\binom{K}{1}$ equações com uma variável independente;
- $\binom{K}{2}$ equações com duas variáveis independentes;
- ...
- $\binom{K}{K}$ equações com todas variáveis independentes;

que somadas totalizam 2^k possíveis equações.

O procedimento computacional que permite a avaliação destas equações é denominado "TODAS POSSÍVEIS REGRESSÕES", o qual pode ser aplicado igualmente bem a dados colineares e não colineares, constituindo uma vantagem em relação aos métodos "stepwise".

Este método fornece ao analista um maior volume de informações com respeito à natureza do relacionamento entre variáveis independentes e a variável resposta. Ao contrário de se ajustar uma única equação, ele fornece outras igualmente, ou até mais promissoras. O critério normalmente utilizado para escolha das equações é a estatística C_p . Aquelas selecionadas são, então, submetidas a análises de resíduos, antes de ser determinada a escolha final.

O grande inconveniente, no entanto, é o excessivo número de equações a serem observadas quando é elevado o número de variáveis independentes ($K > 9$). Este número cresce exponencialmente com K , tornando obrigatória a utilização de um método computacional eficiente.

2.2.2 - Algoritmos

Segundo FURNIVAL (1971), a partir de 1957 inúmeros algoritmos foram desenvolvidos para seleção de variáveis, tendo basicamente, como princípios, minimizar o volume de computações envolvidas e consequentes decréscimos nos tempos de processamentos, sem afetarem, porém, a exatidão dos resultados. Assim, Efroymsen (1960), Hamaker (1962) e Oosterhoff (1963), citados por FURNIVAL (1971), desenvolveram os chamados métodos "stepwise" ou regressão passo a passo.

FURNIVAL e WILSON (1974), GARSIDE (1965) e SCHATZOFF, TSAO e FIENBERG (1968) consideram o método de seleção de todas as possíveis regressões mais informativo, propondo, para tanto, algoritmos que viabilizam os cálculos de maneira mais eficiente. Alguns deles simplesmente fornecem as somas de quadrados residuais, enquanto outros, além delas fornecem as estatísticas usualmente avaliadas em análises de regressão, tais como R^2 , t 's para os parâmetros, etc.

Embora reconhecendo o inconveniente do excessivo número de equações quando é grande o número de variáveis independentes, DANIEL e WOOD (1980) relatam que, utilizando modernos computadores, obtiveram $2^{12} = 4.096$ e $2^{18} = 262.144$ somas de quadrados residuais em menos de 1 segundo e menos de três minutos, respectivamente.

2.2.3 - Métodos de fracionamento

A alternativa para realização de todas as possíveis regressões quando é grande o K, é o que chamamos de fracionamento desta seleção, que consiste na realização de apenas parte dessas equações.

DANIEL e WOOD (1980) e SEBER (1977) relatam um método baseado na observância dos valores de t para os parâmetros da equação completa, ordenados na forma decrescente. A seleção então é feita tomando-se as equações constituídas das variáveis mais importantes em ordem dos maiores valores de t.

GORMAN e TOMAN (1966) sugerem o mesmo algoritmo utilizado por Yates (1937) para fracionamento de tratamentos em experimentos fatoriais da série 2^n , relatados por JOHNSON e LEONE (1977).

HOCKING e LESLIE (1967) propõem um método que permite avaliar para cada grupo de p parâmetros, somente aquelas equações que apresentam menores valores de C_p .

No entanto, DANIEL e WOOD (1980) ressaltam que nenhum destes métodos pode pretender eficiência igual à seleção de todas as possíveis regressões.

2.3 - A Multicolinearidade

2.3.1 - Generalidades

A interpretação de uma equação de regressão múltipla depende implicitamente da suposição de que as variáveis independentes não sejam fortemente correlacionadas.

É usual interpretarmos o coeficiente de regressão como a quantidade de variação sofrida pela variável dependente, quando a variável correspondente a este coeficiente é incrementada em uma unidade, e as demais variáveis mantidas constantes. Este tipo de interpretação para WEISBERG (1980) pode não ter validade se é constatada a existência de acentuadas correlações entre as variáveis independentes ou também denominadas explanatórias. Na maioria dos processos estudados é praticamente impossível modificar o valor de apenas uma variável mantendo as demais inalteradas, tornando assim não válida a usual interpretação do coeficiente de regressão como um efeito marginal.

Duas variáveis são ditas ortogonais se entre elas há uma completa ausência de relacionamento linear. Para CHATTER

JEE e PRICE (1977), desde que não seja acentuada, a condição de não ortogonalidade não chega a afetar significativamente a precisão das estimativas. Esta condição de não ortogonalidade entre variáveis independentes é conhecida como multicolinearidade, embora BELSLEY, KUH e WELSCH (1980) comentem que nenhuma definição tenha sido firmemente estabelecida na literatura. Para estes autores, pode não ser muito fácil a detecção de multicolinearidade, pois não sendo um problema de má formulação do modelo, e sim da própria natureza dos dados, não pode ser constatada por uma análise explorativa dos resíduos.

Recomenda-se bastante cautela sobre todas e quaisquer conclusões substantivas baseadas numa análise de regressão com variáveis colineares. É sempre importante saber, em quaisquer situações, se a multicolinearidade está presente.

2,3.2 - Sintomas

A multicolinearidade está relacionada com a instabilidade na estimação dos coeficientes de regressão. Normalmente as investigações dos problemas provenientes de dados colineares só são possíveis após o modelo já ter sido determinado. No entanto, CHATTERJEE e PRICE (1977) relatam alguns dos indicativos que as vezes são disponíveis por ocasião dos processos de inclusão e exclusão de variáveis ou observações na busca de equações adequadas. São esses os indicativos:

1) Mudanças bruscas nos coeficientes de regressão quando uma variável é incluída ou excluída da equação.

2) Mudanças bruscas nos coeficientes de regressão quando uma observação é descartada, substituída ou adicionada.

Mesmo que a análise dos resíduos tenha dado como satisfatória a equação, a multicolinearidade pode estar presente se:

3) Os sinais algébricos dos coeficientes não estão em conformidade com os esperados a priori.

4) São grandes os erros padrões dos coeficientes de variáveis consideradas importantes.

WEISBERG (1980) sugere o cálculo da estatística denominada "condição número", baseada no cálculo dos autovalores da matriz de correlações, como medida do grau de multicolinearidade. Esta estatística é definida da seguinte forma:

$$\kappa = \{\lambda_1/\lambda_k\}^{0.5} \quad (1)$$

onde λ_1 é o maior autovalor e λ_k o menor, e $k=n-1$ var. ind.

Valores elevados de κ indicam a existência de multicolinearidade. Já CHATTERJEE e PRICE (1977) indicam como medida a soma das recíprocas dos autovalores, dada por:

$$\Psi = \sum_{i=1}^K \frac{1}{\lambda_i} \quad (2)$$

Se $\Psi > 5K$, consideramos as variáveis colineares.

BELSLEY *et alii* (1980) e WEISBERG (1980) abordam com maiores detalhes outros aspectos importantes da multicolinearidade.

2.3.3 - Remédios

Tendo sido constatada a existência de multicolinearidade acentuada, recomenda-se a substituição dos processos usuais de regressão por técnicas mais apuradas, as quais permitirão obter estimativas mais precisas dos parâmetros. HUBER (1981) recomenda a utilização de métodos robustos de estimação, definindo-os como aqueles em que os parâmetros são menos sensíveis às variações dos dados.

CHATTERJEE e PRICE (1977) e WEISBERG (1980) recomendam a aplicação da análise dos componentes principais. MARQUADT e SNEE (1975) e BELSLEY *et alii* (1980) sugerem a utilização do método "ridge" de regressão, e comentam que os seus resultados são muito similares aos obtidos pelos componentes principais. SUNG (1981) sugere uma técnica que consiste na imposição do que chama "ótimas restrições" na estimação do quadrado médio residual. Para este autor, os estimadores $\hat{\beta}$ de mínimos quadrados, sob estas restrições, são praticamente idênticos aos estimadores $\hat{\beta}$ de componentes principais.

2.3.4 - Restrição à utilização dos métodos "stepwise"

Para DANIEL e WOOD (1980), não é recomendável a utilização dos métodos "stepwise" quando é sabida a existência de mul-

colinearidade, pois os resultados podem ser ambíguos. No entanto, se utilizados, a versão BE apresenta melhores resultados do que a FS, devendo, segundo WEISBERG (1980) ser a preferida.

2.4 - Critérios para Seleção de Variáveis

2.4.1 - Generalidades

Após o ajustamento de todas as possíveis regressões, devemos proceder à escolha daquelas mais satisfatórias, utilizando, para tanto, algum critério estatístico. Vários são propostos dentre os quais podemos citar o coeficiente de determinação R^2 , o coeficiente de determinação ajustado \bar{R}^2 , a estatística F, o quadrado médio residual e a estatística C_p como os mais conhecidos e bem caracterizados por HOCKING (1976), SEBER (1977) e WEISBERG (1980). No entanto, autores como GORMAN e TOMAN (1966), HOCKING e LESLIE (1967) e DANIEL e WOOD (1980) recomendam entre esses a estatística C_p de MALLOWS (1964), quando um elevado número de equações de regressão são consideradas.

HOCKING e LESLIE (1967) relatam que antes do seu surgimento os pesquisadores usavam plotar os quadrados médios residuais de cada equação ajustada contra seus respectivos número de parâmetros, o que proporcionava uma melhor visualização daquelas mais promissoras.

2.4.2 - A estatística C_p de MALLOWS (1964)

2.4.2.1 - Definição

A estatística C_p surgiu em 1964 como um aperfeiçoamento implementado por MALLOWS (1964) da usual plotagem dos quadrados médios residuais contra o número de parâmetros. Também conhecida por Erro Quadrado Total Padronizado, ela fornece um total de variações, parte residual e parte devida ao viés do modelo, o que nos leva a defini-la como uma simples função da soma de quadrados residual para cada equação ajustada.

2.4.2.2 - Derivação

A estatística C_p , ou Erro Quadrado Total Padronizado, medida nas N observações, para cada uma das equações ajustadas com p parâmetros, é dada por:

$$\sum_{j=1}^N (v_j - \eta_j)^2 + \sum_{j=1}^N \text{Var} \{Y_{pj}\} \quad ,$$

onde,

$$v_j = v(X_1, X_2, \dots) \quad , \text{ valor esperado pela "equação verdadeira};$$

$$\eta_j = \beta_0 + \sum_{i=1}^K \beta_i X_{ij} \quad , \text{ esperança do valor dado pela equação ajustada.}$$

$v_j - \eta_j = \text{viés}$ na j -ésima observação;
 $p = K + 1$, quando β_0 está presente;
 $p = K$, quando β_0 está ausente.

Definimos, por conveniência,

$$SQVIES_p = \sum_{j=1}^N (v_j - \eta_j)^2$$

e definimos Γ_p o Erro Quadrado Total Padronizado.

Assim,

$$\Gamma_p = \frac{1}{\sigma^2} SQVIES_p + \frac{1}{\sigma^2} \sum_{j=1}^N \text{Var} \{Y_{pj}\}$$

SEBER (1977) demonstra que $\sum_{j=1}^N \text{Var} \{Y_{pj}\} = p\sigma^2$

$$\Gamma_p = \frac{1}{\sigma^2} SQVIES_p + p$$

Para DANIEL e WOOD' (1980) a soma de quadrados residual para uma equação com p parâmetros tem esperança

$$E(SQRES_p) = SQVIES_p + (N-p)\sigma^2$$

e resolvendo para $SQVIES_p$

$$SQVIES_p = E(SQRES_p) - (N-p)\sigma^2$$

e

$$\Gamma_p = \frac{E(SQRES_p)}{\sigma^2} - (N-p) + p$$

$$\Gamma_p = \frac{E(\text{SQRES}_p)}{\sigma^2} - (N-2p)$$

MALLOWS (1973) toma como estimativa de σ^2 (chamemo-la s^2), o quadrado médio residual do modelo completo. Obtemos, então, a estatística C_p , que é uma estimativa de Γ_p .

$$C_p = \frac{1}{s^2} \text{SQRES}_p - (N-2p) \quad (3)$$

Se uma equação com p parâmetros apresenta viés desprezível, $\text{SQVIES}_p \approx 0$, a $E(\text{SQRES}_p) = (N-p)\sigma^2$, e assim

$$E(C_p / \text{VIES}_p = 0) = \frac{1}{\sigma^2} (N-p)\sigma^2 - (N-2p) = p$$

2.4.2.3 - O método gráfico

Se os C_p 's das equações são plotados contra seus respectivos números de parâmetros, aquelas que apresentam menor viés tendem a se agrupar sobre a linha $C_p = p$, enquanto as com maior viés tenderão a localizar-se bem acima desta linha.

Como já foi relatado, o quadrado médio residual do modelo completo é frequentemente utilizado como estimativa de σ^2 , o que força $C_p = K+1$ para a equação completa, e também a assumirmos nenhum viés para esta equação.

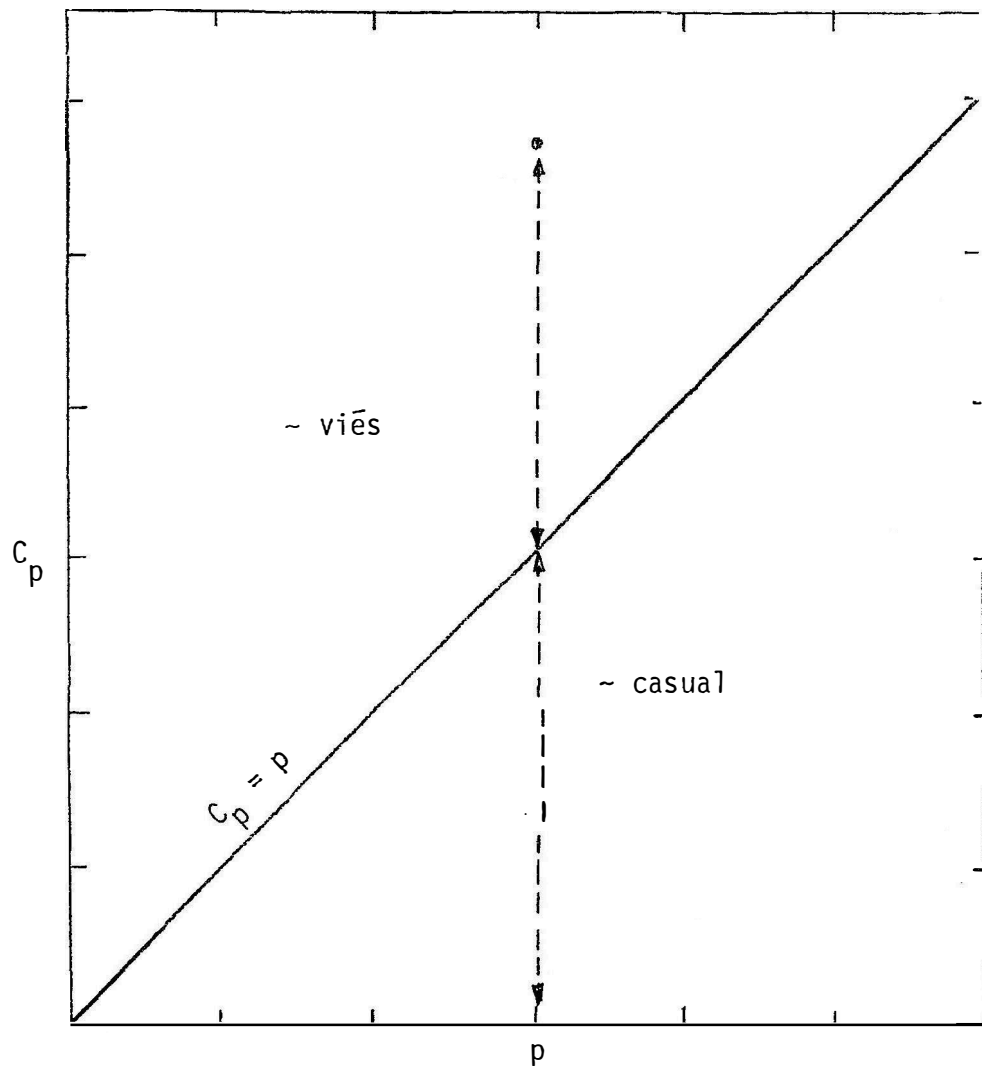


Figura 1 - Gráfico dos C_p 's contra respectivos p 's.

2.5 - Análise de Resíduos

2.5.1 - Generalidades

A análise de resíduos é uma técnica relativamente simples, mas de grande importância na constatação de distúrbios em modelos de análises de regressão.

DRAPER e SMITH (1966) definem os resíduos como as N diferenças do tipo $e_i = Y_i - \hat{Y}_i$, $i=1,2,\dots,N$, onde Y_i é uma observação e \hat{Y}_i o valor estimado correspondente, obtido através da equação ajustada. São também considerados como a quantidade de variação não explicada pela equação.

AFIFI e AZEN (1979) ressaltam que sua grande importância é notada quando utilizada no estudo de violações das hipóteses básicas de independência, variância constante e normalidade dos erros.

SEBER (1977) lembra que este tipo de análise se torna plenamente viável, principalmente pela utilização do computador, sua melhor ferramenta.

Para WEISBERG (1980), muitas vezes, simples diagramas de dispersão dos resíduos são suficientes para afirmarmos se um dado modelo é adequado ou não.

2.5.2 - Métodos gráficos

Normalmente o procedimento mais usual em análises de resíduos é o estudo gráfico. DRAPER e SMITH (1966) e MOSTELLER e TUKEY (1977) apresentam diferentes maneiras para se estudar graficamente os resíduos.

ANSCOMBE (1973) construiu quatro diferentes conjuntos de pares de dados, de maneira que, submetidos a análise de regressão linear simples, apresentam os mesmos valores para as usuais estatísticas \hat{b}_0 , \hat{b}_1 , R^2 , t , s^2 , em contraste com os diagramas de dispersão dos resíduos que se apresentam totalmente diferentes e, na maioria, acentuadamente tendenciosos, o que serve para demonstrar que apenas valores elevados do coeficiente de determinação, assim como valores de t para os parâmetros altamente significativos, não são suficientes para se considerar adequado um ajustamento.

Para SEBER (1977), MONTELLER e TUKEY(1977) e WEISBERG (1980), maior ênfase deve ser dada ao estudo dos resíduos padronizados contra \bar{Y}_j .

Para DRAPER e SMITH (1966) o resíduo padronizado é dado por

$$r_i = \frac{e_i}{s} \quad (4)$$

onde s é desvio padrão residual da equação ajustada. Se o modelo é correto e é assumida a normalidade dos resíduos, então 95% dos r_i 's deverão situar-se entre -2 e +2. Quando plotados contra \bar{Y}_j ou X_{ij} , $j = 1, 2, \dots, K$, os r_i 's não devem exibir nenhuma tendência em torno do

seu valor médio $\bar{r}_i = 0$. Para ser considerado adequado o modelo, eles deverão distribuir-se aleatoriamente na forma de uma faixa horizontal em torno de zero. Se são constatados outros tipos de tendências, ou pontos discrepantes dos demais ("outliers" ou pontos cujo $|r_i| > 2.5$), o modelo, ou parte dele, pode não ser adequado.

A Figura 2 apresenta as tendências mais usualmente observadas nos diagramas de dispersões.

2.5.3 - Estatísticas para exame dos resíduos

DRAPER e SMITH (1966) abordam apenas superficialmente algumas estatísticas para avaliações numéricas de distúrbios em resíduos de regressão, alegando que, na maioria das situações práticas, um exame gráfico dos mesmos é suficientemente informativo para revelar violações nos modelos e indicar uma ação corretiva. Eles abordam, para a avaliação da autocorrelação dos resíduos, a aplicação do Teste dos Chorrilhos, do campo não-paramétrico, citado por HOFFMANN e VIEIRA (1977) e bem detalhado por LINDGRÉN (1976) e FIGUEIREDO (1977), e que se baseia no estudo da distribuição dos sinais (+) e (-) dos r_i s em torno de zero, normalmente ordenados em relação a \bar{Y}_i .

Anscombe (1973), citado por SEBER (1977) e DRAPER e SMITH (1966), propõem, para verificação da homocedasticidade, um teste que consiste basicamente no ajustamento de uma equação de regressão linear simples, onde \bar{Y}_i é considerada a variável independente e

e_j^2 a dependente. A homocedasticidade é aceita se, pelo teste t, o coeficiente b_1 for estatisticamente igual a zero.

SEBER (1977) sugere várias estatísticas apropriadas para teste de normalidade dos resíduos, embora DANIEL e WOOD (1980) achem mais recomendável a utilização do papel normal de probabilidades.

CAMPOS (1979) relata a utilização do teste de Lilliefors para normalidade, quando a média e a variância da amostra não são conhecidas e, sim, estimadas pelos dados.

HAWKINS (1982) propõe um teste de aplicação específica a análises multivariadas para normalidade e homocedasticidade.

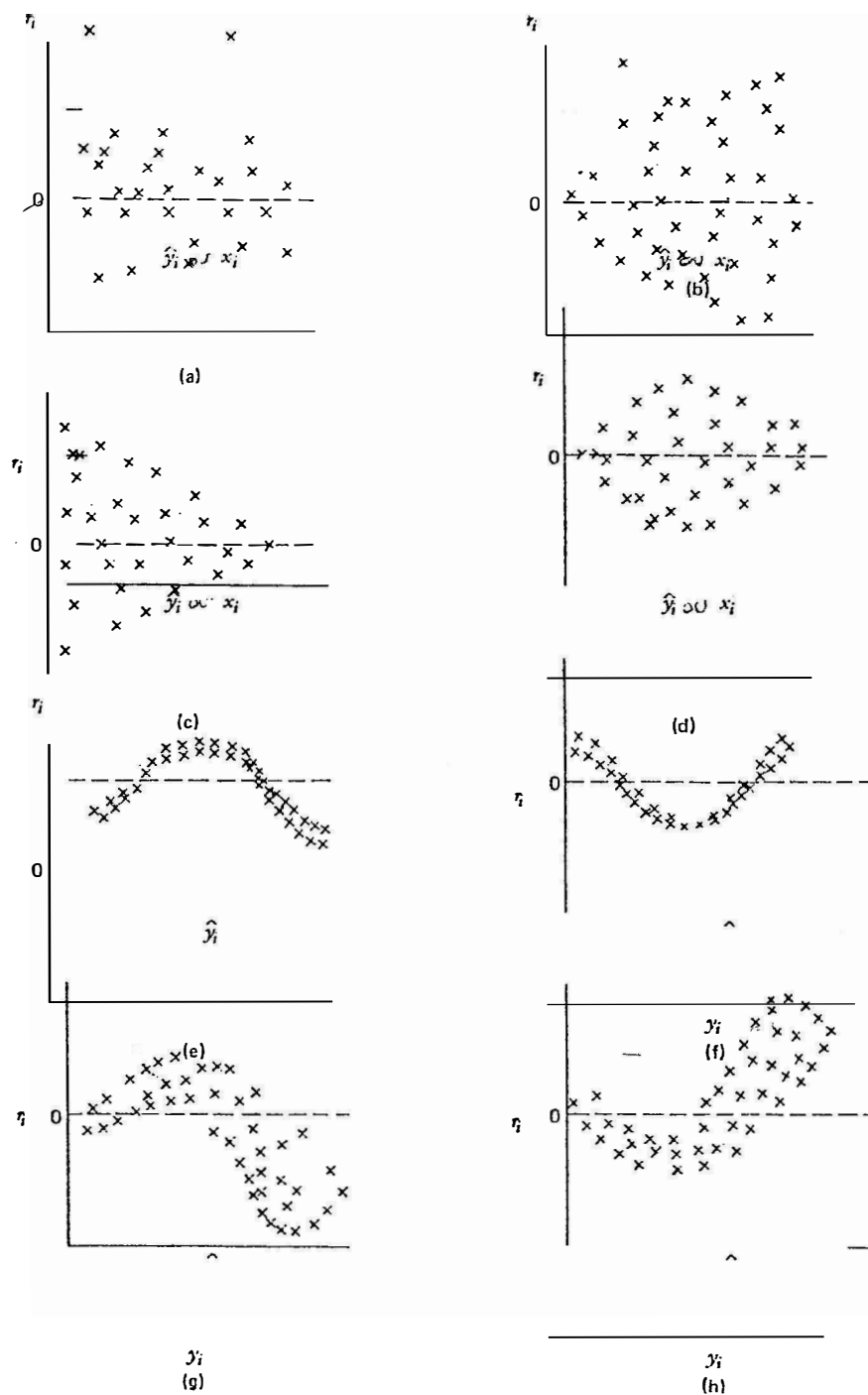


Figura 2 - (a) Faixa horizontal indicando bom ajustamento. (b) Megafone aberto à direita, indicando aumento da variância com \hat{Y}_i ou X_i . (c) Megafone aberto à esquerda, indicando diminuição da variância com \hat{Y}_i ou X_i . (d) Duplo formato de arco, indicando maior variância para valores médios de \hat{Y}_i ou X_i . (e) e (f) indicação de não linearidade do modelo. (g) e (h) Indicação de não linearidade e variância não constante do modelo.

3. MATERIAL E MÉTODOS

3.1 - Material

Os dados utilizados no presente trabalho são provenientes de um experimento conduzido em Curitiba, na câmara de condições controladas do Departamento de Zoologia da Universidade Federal do Paraná, pelo Eng^o-Agr^o José Claret Matioli, os quais gentilmente nos foram concedidos.

Foi instalado um experimento fatorial 4x3x3 com três repetições em delineamento inteiramente casualizado. Cada parcela foi constituída por um frasco de vidro com uma tampa perfurada, cujos orifícios eram vedados por uma tela fina que impedia a fuga dos insetos e permitia as trocas gasosas. Os tratamentos consistiam na infestação artificial em quatro níveis (0, 5, 10 e 20 casais) de *Sí-*

tophilus oryzae (Linnê, 1763), em três cultivares de milho (C-111, FLINT e PIRANÃO) e em três períodos de armazenamento (60, 105 e 150 dias). No final de cada período tomavam-se ao acaso para cada cultivar 12 frascos (4 níveis x 3 repetições), determinando-se as seguintes variáveis:

NINS= X_1 =A: Raiz quadrada do número de insetos + 0,5;

TARM= X_2 =B: Produto do tempo de armazenamento por:

0 se a parcela não foi infestada,

1 se a parcela foi infestada;

CARB= X_3 =C: Teor de carboidratos (%);

UMID= X_4 =D: Teor de umidade (%);

MASE= X_5 =E: Teor de matéria seca (%);

PROT= X_6 =F: Teor de proteínas (%);

OLEO= X_7 =G: Teor de ácidos graxos (%);

IACD= X_8 =H: Índice de acidez dos ácidos graxos (%);

PPES=Y: Arco seno da raiz quadrada da porcentagem de perdas de peso.

Dos dados obtidos, somente as variáveis NINS e PPES apresentaram necessidade de sofrerem transformações para normalização, sendo aplicadas aquelas que melhores resultados apresentaram.

Para a variável TARM, a transformação dicotômica aplicada justificou-se pelo fato de que as parcelas não infestadas apresentam o comportamento no tempo igual ao das parcelas avaliadas

no dia da instalação do experimento, ou seja, aos zero dias.

Os dados são apresentados nas Tabelas 1, 2 e 3.

Tabela 1 - Dados de avaliações de danos em grãos de milho C-111.

OB	NINS	TARM	CARB	UMID	MASE	PROT	OLEO	IACD	PPES
001	0,707	0,0	72,086	12,391	0,987	10,005	4,530	24,336	4,360
002	0,707	0,0	72,865	12,607	0,951	8,936	4,641	26,212	5,437
003	0,707	0,0	72,568	12,529	0,951	9,629	4,322	22,630	5,425
004	6,819	60,0	70,977	12,949	1,031	10,469	4,572	22,113	9,418
005	6,519	60,0	72,167	12,581	0,999	9,777	4,474	28,147	8,148
006	4,743	60,0	72,195	12,615	1,105	9,495	4,588	27,263	7,649
007	12,786	60,0	71,073	13,142	1,041	9,909	4,735	44,133	15,370
008	9,823	60,0	72,216	12,979	1,001	9,035	4,767	36,757	12,485
009	13,472	60,0	72,634	13,039	1,107	8,576	4,644	40,565	14,940
010	13,209	60,0	71,390	13,005	1,040	9,673	4,891	42,461	17,276
011	11,379	60,0	71,801	13,280	1,030	9,027	4,861	40,492	14,586
012	12,980	60,0	71,956	13,048	1,035	9,341	4,620	41,236	16,118
013	0,707	0,0	72,680	12,951	1,155	8,792	4,420	31,440	6,812
014	0,707	0,0	72,867	12,835	1,224	8,520	4,552	33,835	6,632
015	0,707	0,0	72,400	13,024	1,235	8,897	4,444	29,245	7,927
016	8,031	105,0	70,700	13,371	1,287	9,897	4,744	42,660	13,644
017	11,022	105,0	71,325	13,498	1,281	9,323	4,571	43,550	17,484
018	6,123	105,0	71,334	13,206	1,209	9,660	4,590	36,945	9,842
019	14,983	105,0	69,840	13,556	1,393	10,273	4,938	67,534	22,991
020	17,762	105,0	69,274	13,659	1,340	10,856	4,870	97,399	25,814
021	18,207	105,0	69,728	13,457	1,411	10,221	5,182	93,434	25,718
022	14,747	105,0	70,049	13,118	1,326	10,346	5,161	56,625	22,640
023	22,726	105,0	70,028	13,412	1,412	10,473	4,674	148,556	31,502
024	21,737	105,0	68,504	13,951	1,445	11,883	4,215	172,521	30,361
025	0,707	0,0	72,932	12,470	1,210	9,147	4,240	15,842	4,009
026	0,707	0,0	72,225	12,777	1,146	9,456	4,395	15,647	4,887
027	0,707	0,0	72,872	12,842	1,207	8,680	4,397	15,934	5,239
028	13,693	150,0	70,312	13,898	1,226	9,734	4,628	37,867	23,085
029	15,475	150,0	69,680	13,428	1,356	10,740	4,795	36,988	24,806
030	14,300	150,0	71,289	13,503	1,228	9,199	4,780	34,643	22,747
031	13,583	150,0	69,964	13,953	2,418	9,892	3,773	41,520	22,317
032	14,439	150,0	69,531	13,587	2,495	9,588	4,798	35,405	24,088
033	17,930	150,0	70,322	13,241	1,939	9,849	4,648	43,233	30,147
034	19,634	150,0	69,936	13,345	2,307	10,195	4,217	71,769	31,096
035	16,688	150,0	71,369	13,000	1,961	9,023	4,646	38,676	24,646
036	19,912	150,0	70,026	12,934	2,472	10,377	4,189	66,272	30,432

Tabela 2 - Dados de avaliações de danos em grãos de milho FLINT.

OB	NINS	TARM	CARB	UMID	MASE	PROT	OLEO	IACD	PPES
001	0,707	0,0	68,457	11,953	0,958	13,810	4,819	52,665	4,314
002	0,707	0,0	68,382	12,183	1,007	13,689	4,736	46,485	5,367
003	0,707	0,0	68,400	12,129	0,910	13,740	4,819	43,469	4,592
004	4,301	60,0	67,695	12,294	1,263	14,010	4,735	49,628	6,398
005	7,516	60,0	67,152	12,645	1,309	14,119	4,773	45,535	10,861
006	7,382	60,0	66,999	12,810	1,283	14,138	4,767	49,102	9,390
007	11,811	60,0	66,246	12,924	1,353	14,393	5,082	73,487	11,683
008	8,689	60,0	67,148	12,495	1,333	14,012	5,010	67,999	10,216
009	9,772	60,0	67,616	12,974	1,218	13,660	4,529	72,560	10,504
010	13,509	60,0	66,833	13,258	1,026	13,645	5,235	76,588	14,998
011	11,853	60,0	67,052	12,703	1,104	14,082	5,057	51,152	12,828
012	10,653	60,0	67,379	12,692	1,327	13,651	4,949	59,561	13,156
013	0,707	0,0	67,103	13,029	1,808	13,435	4,622	31,989	8,177
014	0,707	0,0	67,544	12,759	1,592	13,470	4,633	32,437	7,321
015	0,707	0,0	67,285	12,690	1,602	13,633	4,787	31,176	7,636
016	12,469	105,0	66,420	13,138	1,768	13,734	4,936	49,933	16,674
017	9,354	105,0	65,472	13,516	1,637	14,442	4,931	39,769	15,150
018	5,244	105,0	66,463	13,245	1,326	14,059	4,905	47,242	10,488
019	13,729	105,0	66,817	13,061	1,668	13,675	4,777	78,162	16,882
020	8,514	105,0	65,875	13,042	1,914	14,350	4,817	81,616	11,922
021	11,510	105,0	66,094	13,540	2,254	13,418	4,732	86,797	16,342
022	13,838	105,0	65,098	13,521	2,048	14,530	4,801	57,121	17,532
023	12,629	105,0	66,961	13,142	1,349	13,767	4,779	53,729	15,808
024	14,017	105,0	66,134	13,219	1,569	14,451	4,624	55,329	16,647
025	0,707	0,0	68,095	12,979	1,497	13,071	4,356	26,710	6,447
026	0,707	0,0	66,437	13,066	1,460	14,253	4,782	22,299	7,156
027	0,707	0,0	67,919	13,203	1,462	12,968	4,445	21,305	7,606
028	14,815	150,0	63,292	14,523	2,387	14,951	4,844	42,634	22,904
029	9,513	150,0	64,680	13,978	1,592	14,762	4,729	31,535	16,041
030	14,089	150,0	64,384	14,040	1,734	14,894	4,946	35,211	21,393
031	13,656	150,0	65,650	14,069	1,435	14,085	4,759	47,302	23,779
032	13,133	150,0	66,222	13,689	1,186	14,039	4,861	40,428	19,382
033	16,896	150,0	64,783	14,049	1,421	15,028	4,716	49,292	26,115
034	16,263	150,0	63,391	14,422	1,139	16,703	4,343	64,506	24,675
035	15,732	150,0	64,398	13,752	1,217	15,662	4,969	43,935	22,855
036	21,459	150,0	62,353	15,646	1,434	18,235	2,330	115,699	32,691

Tabela 3 - Dados de avaliações de danos em grãos de milho PIRANAO.

OB	NINS	TARM	CARB	UMID	MASE	PROT	OLEO	IACD	PPES
001	0,707	0,0	69,498	12,095	1,036	12,849	4,520	20,684	4,761
002	0,707	0,0	69,014	12,124	1,000	13,068	4,791	18,622	4,131
003	0,707	0,0	68,621	12,112	0,850	13,584	4,830	13,974	4,360
004	7,176	60,0	67,591	12,509	1,028	14,067	4,801	13,759	10,167
005	11,423	60,0	66,703	13,057	1,076	14,049	5,112	17,155	12,762
006	8,573	60,0	67,670	12,649	1,157	13,895	4,628	15,751	10,120
007	10,793	60,0	67,106	12,831	1,271	13,910	4,879	17,238	11,394
008	10,024	60,0	67,667	12,835	1,201	13,553	4,743	16,273	11,762
009	7,314	60,0	67,891	12,126	1,194	14,133	4,653	15,453	8,730
010	9,974	60,0	69,535	12,981	1,182	12,807	4,493	13,039	11,155
011	10,464	60,0	67,635	12,489	1,254	14,087	4,533	15,607	11,269
012	9,924	60,0	66,860	12,969	1,301	14,195	4,672	20,872	11,955
013	0,707	0,0	68,479	12,716	1,306	13,175	4,321	10,775	6,971
014	0,707	0,0	68,663	12,721	1,408	13,008	4,296	9,189	6,802
015	0,707	0,0	68,937	12,648	1,362	12,939	4,112	8,624	7,158
016	9,513	105,0	68,593	13,031	1,311	12,531	4,541	20,206	14,485
017	13,583	105,0	68,719	13,003	1,379	12,420	4,477	20,411	15,435
018	11,379	105,0	68,932	12,725	1,337	12,489	4,514	20,352	17,299
019	16,598	105,0	67,298	13,071	1,344	13,597	4,687	31,702	20,418
020	14,781	105,0	67,670	12,875	1,351	13,553	4,548	21,293	17,248
021	10,977	105,0	68,197	12,951	1,369	13,237	4,244	15,885	12,644
022	16,263	105,0	66,351	13,090	1,427	14,610	4,552	33,712	19,985
023	12,708	105,0	67,155	13,820	1,386	13,674	3,962	47,854	17,427
024	17,902	105,0	67,300	12,915	1,309	14,082	4,392	21,446	21,564
025	0,707	0,0	68,850	12,780	1,558	12,680	4,129	10,712	5,802
026	0,707	0,0	66,584	12,713	1,949	14,394	4,359	10,222	6,083
027	0,707	0,0	68,251	12,478	1,812	12,952	4,505	9,307	5,724
028	11,247	150,0	67,515	13,258	1,450	13,282	4,491	20,511	18,006
029	14,815	150,0	63,129	14,215	1,716	16,712	4,227	49,043	23,489
030	11,510	150,0	64,194	13,373	1,582	16,271	4,578	32,980	17,133
031	13,209	150,0	66,350	13,594	1,467	14,478	4,109	34,350	21,939
032	16,263	150,0	65,977	13,807	2,092	13,439	4,682	21,882	23,001
033	10,885	150,0	65,310	13,398	3,064	13,863	4,363	26,330	17,022
034	24,259	150,0	63,145	14,825	3,453	15,762	2,812	125,530	32,163
035	17,449	150,0	61,989	14,708	2,039	17,031	4,230	66,520	29,271
036	21,644	150,0	65,631	14,611	3,000	14,329	2,427	86,761	26,660

3.2 - Métodos

3.2.1 - Introdução

Os microcomputadores, atualmente uma das grandes conquistas da ciência da computação, vêm a cada dia ganhando maiores campos de aplicação, sejam elas domésticas, empresariais ou científicas. A evolução sofrida pela microeletrônica, hoje, nos permite armazenar 1 milhão de bytes ou caracteres, em pastilhas de silício de pouquíssimos mm. A grande importância do desenvolvimento de "softwares" para os microcomputadores pode ser constatada por sua constante carência no mercado, em confronto com a crescente oferta de "hardwares".

A atual política de descentralização no processamento dos dados de pesquisa da EMBRAPA se baseia na utilização em grande escala de microcomputadores em suas unidades de pesquisa, espalhadas por todo o território nacional. Uma equipe especializada vem desenvolvendo um pacote de análises estatísticas denominado SAEST, para microcomputadores POLIMAX, o qual é continuamente aperfeiçoado e implementado com novos programas.

Tendo como objetivo enfatizar a viabilidade de utilização de microcomputadores em análises estatísticas mais complexas, desenvolvemos todo nosso trabalho em um microcomputador PET/CBM 3032 de 64 kbytes de memória, sendo 32 kbytes em RAM e 32 Kbytes em ROM.

Foram também utilizadas duas unidades de floppy disk de 175 Kbytes cada, e uma impressora matricial de 80 colunas e velocidade de 80 cps. As linguagens utilizadas para o desenvolvimento dos programas foram BASIC, relatada por OSBORNE e DONAHUE (1980) e ASSEMBLER, relatada por LEVENTHAL (1979), sendo que estas são linguagens interpretadas, conduzindo, portanto, a um maior gasto de tempo no processamento, se comparadas com outras compiladas, embora sua programação seja mais eficiente.

3.2.2 - Ajustamento das equações completas e avaliação da multicolinearidade

Visando fornecer uma idéia global das variáveis em estudo, serão determinadas as equações constituídas por todas as variáveis, onde os seus parâmetros darão uma visualização daquelas realmente importantes, e se existem indícios de multicolinearidade pela observância de pequenos valores de t em sua maioria. Estas equações também fornecem os valores máximos de R^2 e as estimativas de σ^2 dadas por seus quadrados médios residuais, e utilizadas no cálculo dos C_p 's.

Serão avaliados, conforme recomendações de WEISBERG (1980) e CHATTERJEE e PRICE (1977), os autovalores $\lambda_1 \geq \lambda_2 \geq \dots, \lambda_k$, das matrizes de correlações, determinando-se o grau de multicolinearidade pela estatística Ψ proposta por CHATTERJEE e PRICE (1977).

Para a realização desses cálculos, foi desenvolvido o programa AUTOJACOBI, no qual se utiliza o método de Jacobi, relatado por COOLEY e LOHNES (1962).

3.2.3 - O método para avaliação de todas as regressões possíveis

3.2.3.1 - O algoritmo proposto

A escolha do algoritmo adequado está relacionada com o volume de variáveis a serem estudadas. Alguns autores sugerem que até $K=9$ variáveis independentes é interessante a utilização de um algoritmo que permita avaliar todas as regressões possíveis. O proposto por SCHATZOFF *et alii* (1968), o qual utiliza o operador SWEEP de Beaton, é adequado aos nossos propósitos. A utilização do mesmo nos permite avaliar os quadrados médios residuais para todas $2^K - 1$ equações ajustadas, realizando o mínimo de operações SWEEP nos pivôs das matrizes de correlações.

Basicamente, este operador consiste no seguinte:

Dada uma matriz quadrada $A = \{a_{ij}\}$, dizemos que esta matriz sofreu uma operação SWEEP no k -ésimo pivô, resultando numa matriz $B = \{b_{ij}\}$, se:

$$b_{kk} = \frac{1}{a_{kk}},$$

$$b_{ik} = -\frac{a_{ik}}{a_{kk}}, \quad i \neq k$$

$$b_{kj} = \frac{a_{kj}}{a_{kk}}, \quad j \neq k$$

$$b_{ij} = a_{ij} - \frac{a_{ik}a_{kj}}{a_{kk}}, \quad i, j \neq k$$

O SWEEP também apresenta as seguintes características teóricas:

1) O SWEEP é reversível:

Aplicar o operador duas vezes consecutivas no mesmo pivô, corresponde a não ter aplicado nenhuma vez.

2) O SWEEP é comutativo:

Aplicar o SWEEP no r -ésimo e em seguida no s -ésimo pivô, equivale à mesma aplicação na ordem inversa.

Maiores detalhes sobre este operador são relatados por SEBER (1977), SCHATZOFF *et alii* (1968) e GOODNIGHT (1979).

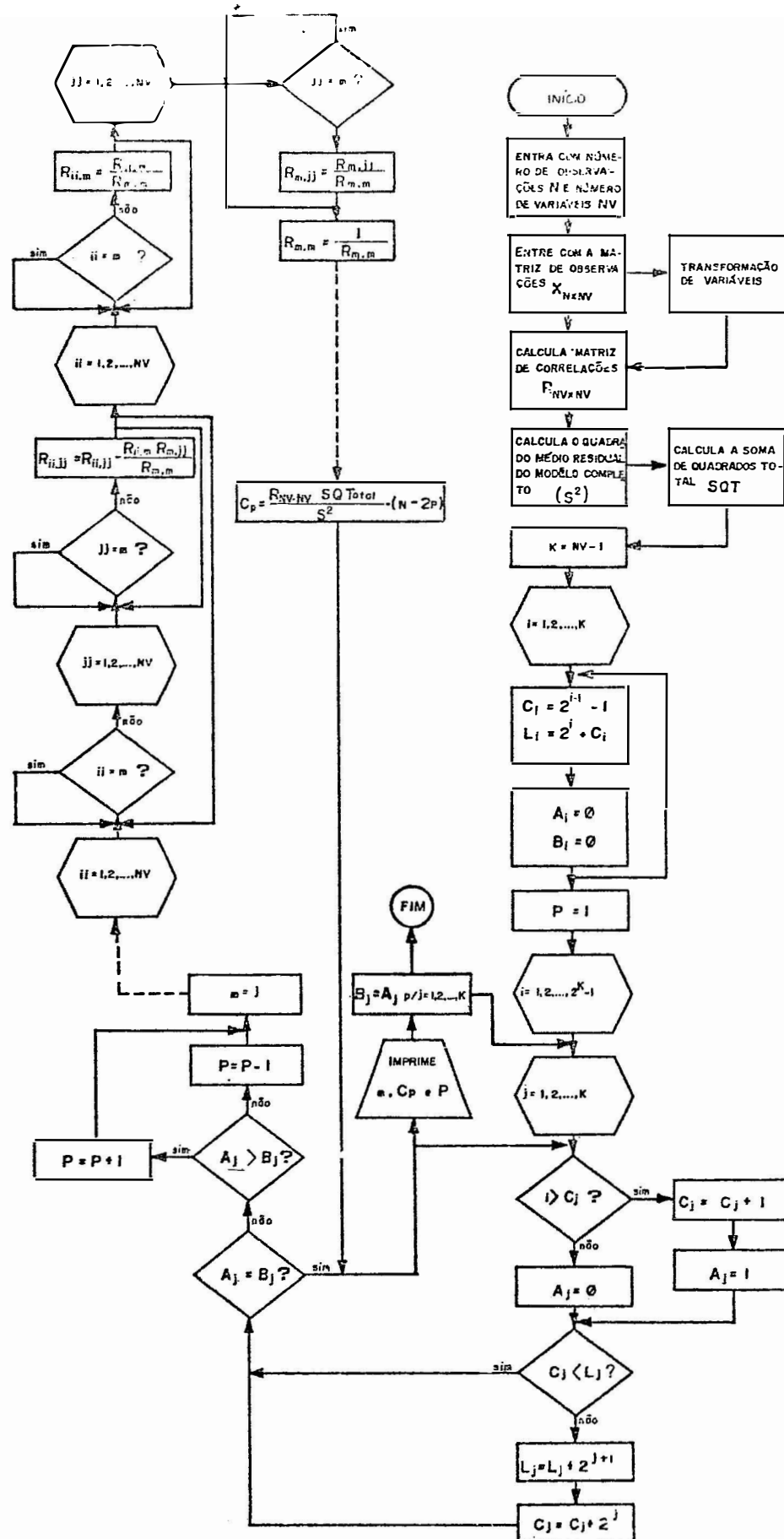


Figura 3 - Algoritmo proposto para cálculo dos C_p 's de todas as possíveis regressões.

A parte principal do algoritmo é a geração da sequência ótima de pivôs, os quais sofrerão operação SWEEP. Esta sequência nos permite avaliar, para um conjunto de K variáveis independentes, as $2^K - 1$ equações de regressão, com um número mínimo de SWEEP's também $2^K - 1$. Baseados nos trabalhos de SCHATZOFF *et alii* (1968), criamos um algoritmo para aplicação em microcomputadores, que permite gerar esta sequência ótima de pivôs. No entanto, não devemos considerá-lo mais eficiente, podendo eventualmente surgirem outros mais bem sucedidos. O Quadro 1 representa o resultado da aplicação deste algoritmo para um exemplo hipotético com $K=5$ variáveis independentes.

Quadro 1 - Sequência ótima de pivôs gerada pelo algoritmo proposto.

Passo	Sequencia otima de pivos	x's 12345	Modelos
1	1	10000	x1, Y
2	2	11000	x1, x2, Y
3	1	01000	x2, Y
4	3	01100	x2, x3, Y
5	1	11100	x1, x2, x3, Y
6	2	10100	x1, x3, Y
7	1	00100	x3, Y
8	4	00110	x3, x4, Y
9	1	10110	x1, x3, x4, Y
10	2	11110	x1, x2, x3, x4, Y
11	1	01110	x2, x3, x4, Y
12	3	01010	x2, x4, Y
13	1	11010	x1, x2, x4, Y
14	2	10010	x1, x4, Y
15	1	00010	x4, Y
16	5	00011	x4, x5, Y
17	1	10011	x1, x4, x5, Y
18	2	11011	x1, x2, x4, x5, Y
19	1	01011	x2, x4, x5, Y
20	3	01111	x2, x3, x4, x5, Y
21	1	11111	x1, x2, x3, x4, x5, Y
22	2	10111	x1, x3, x4, x5, Y
23	1	00111	x3, x4, x5, Y
24	4	00101	x3, x5, Y
25	1	10101	x1, x3, x5, Y
26	2	11101	x1, x2, x3, x5, Y
27	1	01101	x2, x3, x5, Y
28	3	01001	x2, x5, Y
29	1	11001	x1, x2, x5, Y
30	2	10001	x1, x5, Y
31	1	00001	x5, Y

3.2.3.2 - O cálculo dos C_p 's

Para a avaliação de todas as possíveis regressões, foi desenvolvido um programa denominado MALLOW'S. Este programa foi elaborado utilizando o algoritmo anteriormente descrito, o qual simplesmente calcula a soma de quadrados residuais para cada equação, seu respectivo número de parâmetros e o valor da estatística C_p . Classificadas em ordem crescente, as vinte primeiras equações que apresentarem menores valores de C_p serão consideradas como alternativas. A equação para cada cultivar, que apresentar o menor valor de C_p deverá ser a escolhida para ser submetida às análises complementares de avaliação e teste de significância dos parâmetros, assim como seus respectivos intervalos de confiança. Para essas análises complementares foi desenvolvido um programa denominado STEP.COMP, o qual realiza também a aplicação dos testes de chorrilhos e de Lilliefors aos resíduos. A equação escolhida deverá também ser submetida a uma análise gráfica dos resíduos.

3.2.4 - Análise gráfica dos resíduos

Os resíduos da equação selecionada depois de devidamente padronizados, serão estudados graficamente contra as variáveis independentes consideradas relevantes na equação escolhida, e

contra os valores \hat{Y}_i estimados por essa equação, objetivando visualizar distorções no modelo.

DRAPER e SMITH (1966) definem também os resíduos padronizados como desvios normais unitários, os quais têm média zero e variância 1.

$$r_i = \frac{Y_i - \hat{Y}_i}{s} \quad , \quad i=1,2,\dots,N \quad . \quad (5)$$

onde r_i representa o i -ésimo resíduo padronizado, Y_i o i -ésimo Y observado e \hat{Y}_i o i -ésimo Y estimado pela equação selecionada, e s é o desvio padrão residual da equação, dado pela raiz quadrada do quadrado médio residual.

Para a realização destas análises gráficas, desenvolvemos um programa denominado PLOT.RES. Este programa determina que os resíduos padronizados sejam plotados na ordenada do gráfico, cuja amplitude é de -3 a +3, e as variáveis X_i ou \hat{Y}_i consideradas a abscissa do eixo.

3.2.5 - O teste dos chorrilhos para autocorrelação dos resíduos

A utilização deste teste em análises de resíduos de regressão, é bem enfatizada por DRAPER e SMITH (1966). Pertencente ao campo da estatística não-paramétrica, tem como principal aplica-

ção verificar se a hipótese de aleatoriedade de uma amostra não é rejeitada. Aplicado em análise de resíduos de regressão, este teste nos fornece uma medida do grau de autocorrelação destes resíduos, os quais, para este teste, devem ser ordenados em relação a \hat{Y} . ou alternativamente em relação as variáveis independentes. A autocorrelação é investigada através da distribuição dos sinais (+) e (-) em torno do resíduo médio $\bar{e}=0$.

Para melhor visualização, considere a sequência abaixo de sinais, como uma sequência hipotética de $N=14$ sinais de resíduos ordenados em relação a \hat{Y}_i .

(+ +)(-)(- - - -)(+ +)(-)(+ + + +) .

Um chorrilho é então definido como uma sucessão de sinais idênticos, representados pelos grupos de sinais dentro dos parênteses. Neste exemplo hipotético para $n_1=8$ sinais positivos, e $n_2=N-n_1=6$ sinais negativos, foram contados $r=6$ chorrilhos.

Para várias aplicações, o teste é bilateral, ou seja, tanto um número pequeno quanto a um número grande de chorrilhos, são indicativos de ausência de aleatoriedade. Em se tratando de aplicação em análises de resíduos, será considerada existência de autocorrelação somente para pequenos valores de r .

O número total de chorrilhos representado pela letra r apresenta então a seguinte distribuição de probabilidades.

$$f(r) = \left\{ \begin{array}{l} \frac{2 \binom{n_1 - 1}{\frac{r}{2} - 1} \binom{n_2 - 1}{\frac{r}{2} - 1}}{\binom{n_1 + n_2}{n_1}}, \text{ se } r \text{ é par} \\ \frac{\binom{n_1 - 1}{\frac{r-1}{2}} \binom{n_2 - 1}{\frac{r-3}{2}} + \binom{n_1 - 1}{\frac{r-3}{2}} \binom{n_2 - 1}{\frac{r-1}{2}}}{\binom{n_1 + n_2}{n_1}}, \text{ se } r \text{ é ímpar} \end{array} \right.$$

$n_1 + n_2$.

DRAPER e SMITH (1966) sugerem a aproximação normal quando N for maior do que 20. Para o presente trabalho foi elaborada uma subrotina dentro do programa STEP.COMP, que possibilita a geração da distribuição nula de probabilidades do teste para N até 50, permitindo com isso um cálculo mais exato das probabilidades. Para N maior que 50 a subrotina aplica a aproximação normal, onde

$$z = \frac{r - \mu + 1/2}{\sigma} \quad (6)$$

FIGUEIREDO (1977) demonstra que a média e a variância são definidas por:

$$\mu = \frac{2n_1n_2}{n_1+n_2} + 1 \quad (7)$$

$$\sigma^2 = \frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)} \quad (8)$$

As hipóteses formuladas são:

H_0 : Os resíduos se distribuem aleatoriamente em torno de zero
(os erros não são autocorrelacionados);

H_1 : Os resíduos não se distribuem aleatoriamente em torno de
zero (os erros são autocorrelacionados).

Rejeitamos H_0 ao nível $\alpha = 0,1$, se $P(r \leq r \text{ observado}) < \alpha$.

3.2.6 - O teste de Lilliefors para normalidade dos resíduos

Conforme CAMPOS (1979), este teste foi introduzido por Lilliefors em 1967, como uma modificação do teste de Kolmogorov-Smirnov, possibilitando o seu uso para casos em que a média e a variância não são especificadas, e, sim, estimadas através dos dados da amostra.

Este teste pode ser aplicado sem restrições para pequenas amostras, constituindo uma vantagem em relação ao teste de χ^2 . Uma outra vantagem é a não perda de informações devido ao agru-

pamento de dados, como ocorre no teste de χ^2 , pois os dados são considerados individualmente.

Em se tratando da aplicação deste teste para verificação da normalidade dos resíduos, a média tem valor zero e a variância é dada pelo quadrado médio residual da equação ajustadas, e os dados são os e_i 's estimados. Após a padronização dos e_i 's já anteriormente relatada, é feita a ordenação dos mesmos em ordem crescente. A partir daí consideramos:

$F(r_i)$ = proporção de valores esperados $\leq r_i$; valor gerado pelo programa STEP.COMP;

$S(r_i)$ = proporção de valores obtidos $\leq r_i$; $S(r_i) = \frac{q}{N}$, onde q é o número de valores obtidos a partir dos valores observados.

Conforme CAMPOS (1979)

$$D = \sup_{r_i} |F(r_i) - S(r_i)| \quad (9)$$

onde \sup_{r_i} = supremo em relação a r_i , isto é, a máxima distância vertical entre $F(r_i)$ e $S(r_i)$.

Para a determinação de D , considera-se em cada ponto r_i as diferenças

$$|F(r_i) - S(r_i)| \quad \text{e} \quad |F(r_i) - S(r_{i-1})| \quad ,$$

tomando-se a maior delas.

O teste é bilateral, onde se tem:

H_0 : Os resíduos são considerados normalmente distribuídos;

H_1 : Os resíduos não são normalmente distribuídos.

Rejeitamos H_0 a um nível $\alpha = 0,05$ de probabilidade, quando $D \geq d$, onde os valores de d são consultados numa tabela previamente gravada pelo programa STEP.COMP.

O teste é aproximado, uma vez que os limites da tabela são determinados empiricamente.

3.2.7 - Determinação de modelos pelo método dos componentes principais

Para fins ilustrativos, serão estimados, para os três conjuntos de dados, modelos pelo método dos componentes principais, relatado por MORRISON (1967). O objetivo é verificar se apresentam melhores estimativas das perdas de peso do que aquelas selecionadas pelo método de todas as possíveis regressões.

Para sua aplicação, CHATTERJEE e PRICE (1977) sugerem que as variáveis independentes sejam padronizadas, já que os componentes principais são combinações lineares das mesmas, e estas, por sua vez, medidas em diferentes unidades (cm, gramas, %, etc.), tornando problemática a determinação da unidade resultante. A padro

nização faz com que as variáveis fiquem adimensionais.

A matriz de correlações para as variáveis padronizadas é idêntica à matriz de correlações para as variáveis originais, conseqüentemente, os autovalores $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k$ e seus respectivos autovetores são os mesmos. Os componentes principais são obtidos multiplicando-se os autovetores pela transposta da matriz de variáveis padronizadas, obtendo-se uma nova matriz ortogonal, a qual será submetida a uma análise de regressão pelo método STEPWISE, determinando-se as porcentagens de variação explicada dos componentes e a seleção daqueles mais expressivos.

A principal finalidade destes ajustamentos é verificar se os resíduos dos mesmos se apresentam melhor distribuídos do que dos modelos anteriormente ajustados. Se os resíduos provenientes dos modelos de componentes principais se apresentarem melhores distribuídos, fica constatada a influência negativa da multicolinearidade, caso contrário podemos reafirmar a suposição de que os distúrbios são inerentes da utilização de modelos lineares, não adequados a estes estudos.

4. RESULTADOS E DISCUSSÃO

Os dados utilizados apresentam características que servem bem aos propósitos do presente trabalho, pois algumas das variáveis estudadas apresentam comportamentos já conhecidos.

Os três cultivares de milho estudados apresentam diferenças físicas e químicas entre si, relatadas por MATIOLI (1978), o que nos levou a esperar, a priori, diferentes equações finais para os mesmos.

Os ajustamentos das equações completas foram obtidos pelo programa STEP.COMP e são apresentados nas Tabelas 4, 5 e 6, as quais englobam essencialmente as estatísticas usualmente avaliadas em análises de regressão linear múltipla.

Nota-se que, para os três cultivares, a variável NINS apresentou, pelo teste t a 5%, resultado significativo. Este resul-

tado já era esperado, em virtude da população de insetos ser o fator marcadamente mais influente na determinação das perdas de peso.

Uma outra similaridade entre as equações completas, é a alta porcentagem de variação explicada obtidas pelas mesmas, da ordem de aproximadamente 97%, e também os elevados valores de F constatados nas análises da variância, aparentemente indicando um excelente ajustamento.

Excluída NINS, os cultivares C-111 e PIRANÃO apresentaram valores pequenos de $|t|$ (em torno de 1) para as estimativas dos parâmetros das demais variáveis, o que, segundo CHATTERJEE e PRICE (1977) pode ser um indicativo da existência de multicolinearidade. O cultivar FLINT apresentou um maior número de parâmetros significativos ($|t's| > 2$), indicando um maior número de variáveis associadas a PPES.

Tabela 4 - Estatísticas obtidas para a equação completa referente ao cultivar C-111.

TERMOS	EQUAÇÃO COMPLETA				COEFICIENTE DE CORRELAÇÕES SIMPLES									
	b_j	$s(b_j)$	t_j	I.C _{0,05}	NINS	TARM	CARB	UMID	MASE	PROT	OLEO	IACE	PPES	
Constante	967,38.				NINS									
NINS = X ₁	0,7832	0,1384	5,66*	$b_1 \pm 0,2834$	TARM	0,8416	1,0							
TARM = X ₂	0,0170	0,0169	1,01	$b_2 \pm 0,0347$	CARB	-0,8332	-0,8035	1,0						
CARB = X ₃	-9,9093	9,3443	-1,06	$b_3 \pm 19,1558$	UMID	0,7047	0,7350	-0,7931	1,0					
UMID = X ₄	-9,1341	9,8436	-0,93	$b_4 \pm 20,1794$	MASE	0,5362	0,6743	-0,5729	0,4338	1,0				
MASE = X ₅	-5,0432	9,0245	-0,56	$b_5 \pm 18,5043$	PROT	0,6023	0,4834	-0,8390	0,4672	0,2511	1,0			
PROT = X ₆	-9,7294	9,3880	-1,04	$b_6 \pm 19,2454$	OLEO	0,2358	0,1376	-0,1701	0,1087	-0,3221	0,0403	1,0		
OLEO = X ₇	-7,7266	9,1897	-0,84	$b_7 \pm 18,8389$	IACD	0,7158	0,3897	-0,6953	0,5652	0,2360	0,6953	0,0824	1,0	
IACD = X ₈	0,0356	0,0214	1,66	$b_8 \pm 0,0439$	PPES	0,9664	0,8728	-0,8704	0,7041	0,6553	0,6001	0,1647	0,6941	1,0
N = 36					Média	10,530	78,750	71,198	13,144	1,361	9,691	4,598	47,330	16,669
SQRES = 84,02					D.Padrão	7,063	56,282	1,241	0,413	0,443	0,725	0,278	33,713	9,135
Teste F = 113,95**					C.V.(%)	67,47	71,46	1,74	3,14	32,58	7,47	6,04	71,23	54,80
R ² = 97,12%														
s ² = 3,1119														

* Significativo ao nível de 5%.
Os coeficientes de correlações sublinhados representam correlações acentuadas entre variáveis independentes.

Tabela 5 - Estatísticas obtidas para a equação completa referente ao cultivar FLINT.

TERMOS	EQUAÇÃO COMPLETA				COEFICIENTES DE CORRELAÇÕES SIMPLES									
	b_i	$s(b_i)$	t_i	I.C. 0,5	NINS	TARM	CARB	UMID	MASE	PROT	OLEO	IACD	PPES	
Constante	-1304,75				NINS									
X_1	0,6351	0,1172	5,42*	$b_1 \pm 0,2403$	TARM	0,8795	1,0							
X_2	0,0186	0,0112	1,66	$b_2 \pm 0,0230$	CARB	-0,7700	-0,8261	1,0						
X_3	12,6984	5,7703	2,20*	$b_3 \pm 11,8291$	UMID	0,7278	0,7719	-0,9221	1,0					
X_4	16,2135	5,8578	2,77*	$b_4 \pm 12,0085$	MASE	0,1884	0,2910	-0,4061	0,3808	1,0				
X_5	12,2075	5,7408	2,13*	$b_5 \pm 11,7686$	PROT	0,6497	0,6199	-0,8492	0,7525	-0,0017	1,0			
X_6	12,9234	5,7993	2,22*	$b_6 \pm 11,8886$	OLEO	-0,2100	-0,1272	0,3654	-0,5404	-0,0276	-0,6228	1,0		
X_7	11,6344	5,8324	1,99	$b_7 \pm 11,9564$	IACD	0,5794	0,3425	-0,2988	0,2731	0,0248	0,4308	-0,4000	1,0	
X_8	-0,0418	0,0194	2,15*	$b_8 \pm 0,0398$	PPES	0,9172	0,8902	-0,8910	0,9029	0,2668	0,7505	-0,3812	0,4099	1,0
N = 36					Média	9,131	78,750	66,339	13,233	1,461	14,237	4,720	52,066	14,054
SQRES = 47,37					D. Padrão	5,895	56,282	1,501	0,748	0,339	0,988	0,450	19,901	6,881
Teste F = 114,69**					C.V.(%)	64,56	71,46	2,26	5,65	23,19	6,94	9,54	38,22	48,96
$R^2 = 97,14\%$														
$s^2 = 1,7545$														

* Significativo ao nível de 5%.

Tabela 6 - Estatísticas obtidas para a equação completa referente ao cultivar PIRANAO.

TERMOS	EQUAÇÃO COMPLETA				COEFICIENTES DE CORRELAÇÕES SIMPLES								
	b_i	$s(b_i)$	t_i	I.C. ₉₅	NINS	TARM	CARB	UMID	MASE	PROT	OLEO	IACD	PPES
Constante	158,00				NINS								
NINS = X ₁	0,5017	0,1102	4,55*	$b_1 \pm 0,2259$	TARM	0,8807	1,0						
TARM = X ₂	0,0280	0,0123	2,28*	$b_2 \pm 0,0252$	CARB	-0,6357	-0,6761	1,0					
CARB = X ₃	-1,9852	1,7219	-1,15	$b_3 \pm 3,5299$	UMID	0,7653	0,7679	-0,8171	1,0				
UMID = X ₄	1,4597	1,8144	0,81	$b_4 \pm 3,7125$	MASE	0,4911	0,5018	-0,6394	0,7334	1,0			
MASE = X ₅	-2,1029	1,9838	-1,06	$b_5 \pm 4,0668$	PROT	0,4906	0,5030	-0,9231	0,6267	0,3897	1,0		
PROT = X ₆	-1,9288	1,8108	-1,06	$b_6 \pm 3,7121$	OLEO	-0,4154	-0,3328	0,3710	-0,6666	-0,7473	-0,2496	1,0	
OLEO = X ₇	-1,6611	2,1125	-0,79	$b_7 \pm 4,3306$	IACD	0,7098	0,5945	-0,7168	0,8401	0,7342	0,6027	-0,7833	1,0
IACD = X ₈	0,0274	0,0330	0,83	$b_8 \pm 0,0677$	PPES	0,9367	0,9047	-0,7619	0,8904	0,6125	0,5909	-0,5173	0,7913
N = 36					Média	9,917	78,750	67,195	13,058	1,528	13,853	4,395	26,501
SQRES = 69,96					D. Padrão	6,459	56,282	1,816	0,694	0,575	1,114	0,502	23,581
Teste F = 86,95**					C.V.(%)	65,12	71,46	2,70	5,31	37,64	8,04	11,42	88,98
R ² = 96,26%													50,99
s ² = 2,5912													

* Significativo ao nível de 5%.

Das matrizes de correlações encontradas nas Tabelas 4, 5 e 6, as partes referentes às correlações entre as variáveis independentes, foram submetidas ao programa AUTO-JACOBI para determinação dos autovalores, os quais são apresentados na Tabela 7.

Tabela 7 - Autovalores obtidos para as matrizes de correlações.

λ_i	C-111	FLINT	PIRANÃO
1	4,76498	4,65292	5,45582
2	1,29758	1,28269	1,13609
3	0,91669	0,93397	0,76444
4	0,42791	0,76141	0,32163
5	0,35884	0,21394	0,14597
6	0,18870	0,08873	0,10576
7	0,04486	0,06595	0,06584
8	0,00040	0,00036	0,00442
$\Psi = \sum_{i=1} 1/\lambda_i$	2.534,78	2.812,26	263,22

Valor limite: $\Psi = 5(K=8) = 40$.

As somas das recíprocas dos autovalores excederam em muito o valor limite $\Psi = 40$, indicando a existência de acentuada multicolinearidade. A obtenção de pequenos autovalores indicam que algumas variáveis não são relevantes por estarem acentuadamente correlacionadas com outras. Nestas circunstâncias, recomendamos a não utilização dos métodos "stepwise" de regressão para seleção de variáveis, conforme sugestão de DANIEL e WOOD (1980). A avaliação de to-

das as possíveis regressões é mais adequada por não sofrer influência da multicolinearidade.

5.3 - Obtenção de todas as Possíveis Regressões

Os conjuntos de dados foram analisados através do programa MALLOW'S, e as 256 equações ajustadas para cada cultivar se encontram nas Tabelas 8, 9 e 10. Nestas tabelas as setas situadas à direita das letras representativas das variáveis indicam que estas equações apresentam valores C_p 's \leq p's.

A Tabela 11 apresenta, em ordem crescente de C_p 's, as vinte equações alternativas escolhidas para cada cultivar.

Notamos nesta tabela, que em todas as 60 equações, a variável NINS, representada pela letra A, está presente, confirmando definitivamente a importância desta variável, antecipadamente prevista.

Para o cultivar C-111 observa-se que em 19 das 20 equações a variável MASE, representada pela letra E, está presente, assim como a variável UMID, representada pela letra D, está presente em todas as 40 equações referentes aos cultivares FLINT e PIRANÃO, nos indicando a relevância destas variáveis. Coincidentemente, as equações que apresentaram menores C_p 's para cada cultivar são constituídas por três variáveis.

Tabela 8 - Todas as possíveis regressões fornecidas pelo programa MALLOW'S, para o cultivar C-111

Variaáveis	R ²	Cp	Variaáveis	R ²	Cp	Variaáveis	R ²	Cp
Nenhuma	1	904.66	IA	2	29.98	IAB	3	20.60
IA	2	191.55	IAC	3	117.80	IABC	4	15.76
IAC	3	18.93	IC	2	195.46	ICD	3	191.20
IACD	4	20.55	IACD	5	17.71	IACD	4	119.79
IAD	3	173.72	IAD	4	20.64	IAD	3	25.49
IAD	2	392.50	IAD	3	265.31	IAD	4	4.84
IADDE	5	6.74	IADDE	4	164.53	IADDE	5	116.87
IADDE	6	6.60	IADDE+	5	4.65	ICDE	4	157.63
ICDE	3	163.16	ICDE+	4	3.67	ICDE	5	5.42
ICDE	4	115.13	ICDE	3	185.89	IADDE	4	8.06
IADDE	3	7.22	IE	2	503.51	IEF	3	315.44
IADDE	4	7.71	IADDE	5	8.54	IEF	4	144.30
IADDE	5	113.40	IADDE	6	6.02	ICDE+	5	4.02
IADDE	4	132.79	ICDE	5	140.10	ICDEF	6	6.01
ICDEF	7	8.01	ICDEF	6	114.15	IADDE	5	132.43
ICDEF	6	7.61	IADDE	5	5.73	IADDE	4	201.09
IADDE	3	316.81	IADDE	4	27.24	IADDE	5	22.10
IADDE	4	145.52	IADDE	5	112.52	IADDE	6	12.37
IADDE	5	10.69	ICDF	4	140.15	ICF	3	143.68
IADDE	4	10.67	IADDE	5	12.17	ICF	4	113.20
IADDE	3	154.66	IADDE	4	21.83	IADDE	3	31.50
IADDE	2	568.57	IADDE	3	562.00	IADDE	4	29.83
IADDE	5	21.57	IADDE	4	154.20	ICDFG	5	115.19
IADDE	6	9.30	ICDFG	5	7.35	ICFG	4	144.87
IADDE	5	140.18	ICDFG+	6	5.79	IADDE	7	7.78
IADDE	6	114.29	IADDE	5	145.23	IADDE	6	21.96
IADDE	5	26.22	IADDE	4	311.51	IADDE	5	141.00
IADDE	6	6.24	IADDE	7	8.23	IADDE	6	113.35
IADDE	7	115.29	ICDEFG	8	9.76	ICDEFG	7	7.77
ICDEFG	6	142.09	ICDEFG	5	140.10	ICDEFG+	6	5.98
ICDEFG	7	7.98	ICDEFG	6	113.79	IEFG	5	127.03
IADDE	6	9.59	IADDE	5	8.31	IEFG	4	208.61
IEG	3	357.52	IADDE	4	8.46	IADDE	5	9.62
IEG	4	175.03	ICDEG	5	112.50	IADDE	6	7.17
ICDEG	5	5.33	ICDEG	4	147.80	ICDEG	5	140.37
IADDE	6	6.13	IADDE	7	8.13	ICDEG	6	113.63
IADDE	5	150.98	IADDE	6	7.92	IADDE	5	5.93
IEG	4	203.36	IADDE	3	387.75	IADDE	4	24.30
IADDE	5	20.24	IADDE	4	173.95	ICDGG	5	121.53
IADDE	6	17.23	IADDE	5	19.22	ICDGG	4	192.80
ICG	3	197.19	ICG	4	17.64	IADDE	5	15.25
ICG	4	119.54	ICG	3	191.65	IADDE	4	20.01
IADDE	3	28.02	IG	2	881.18	IGH	3	445.47
IADDE	4	29.99	IADDE	5	17.56	IADDE	4	55.88
ICGH	5	57.54	IADDE	6	16.76	ICGH	5	18.48
ICGH	4	184.60	ICDGH	5	189.76	IADDE	6	20.09
IADDE	7	18.76	ICDGH	6	58.29	IADDE	5	57.07
IADDE	6	19.35	IADDE	5	26.12	IDGH	4	286.15
IADDE	5	101.18	IADDE	6	5.74	IADDE	7	6.22
IADDE	6	38.45	ICDEGH	7	49.19	IADDE	8	8.07
ICDEGH	7	7.07	ICDEGH	6	96.30	ICGH	5	103.61
IADDE	6	6.47	IADDE	7	6.64	ICGH	6	38.19
IADDE	5	36.56	IADDE	6	5.04	IADDE	5	7.30
IEGH	4	130.34	IEFGH	5	123.95	IADDE	6	8.68
IADDE	7	7.06	IEFGH	6	38.16	ICDEFGH	7	49.15
IADDE	8	7.86	ICDEFGH+	7	6.78	ICDEFGH	6	94.30
ICDEFGH	7	90.81	ICDEFGH	8	8.01	IADDE	9	9.00
ICDEFGH	8	39.03	IADDE	7	46.05	IADDE	8	8.12
IADDE	7	7.18	IADDE	6	97.20	IADDE	5	275.89
IADDE	6	27.33	IADDE	7	21.15	IADDE	6	56.60
ICDFGH	7	38.41	IADDE	8	7.31	ICDFGH+	7	6.28
ICDFGH	6	91.53	ICFGH	5	109.04	IADDE	6	9.09
IADDE	7	10.54	ICFGH	6	50.43	IADDE	5	55.80
IADDE	6	19.29	IADDE	5	31.67	IFGH	4	421.58
IFH	3	431.31	IFH	4	33.43	IADDE	5	18.45
IFH	4	54.63	ICDFH	5	48.47	IADDE	6	12.16
IADDE	5	12.25	ICDFH	4	107.90	ICDFH	5	92.89
IADDE	6	11.23	IADDE	7	19.16	ICDFH	6	36.80
IADDE	5	55.41	IADDE	6	20.36	IADDE	5	28.98
IADDE	4	279.91	IADDE	5	155.44	IADDE	6	7.27
IADDE	7	7.94	IADDE	6	48.65	ICDEFH	7	38.58
IADDE	8	7.70	IADDE	7	6.73	ICDEFH	6	92.29
ICDFH+	5	98.15	ICDFH+	6	4.78	IADDE	7	5.98
ICDFH	6	45.57	IEFH	5	47.14	IADDE	6	6.79
IADDE	5	8.85	IEFH	4	208.61	IEH	3	216.23
IADDE	4	7.23	IADDE	5	4.79	IEH	4	46.42
ICDEH	5	48.35	IADDE	6	5.77	IADDE	5	5.30
ICDEH	4	133.56	ICDEH	5	128.66	IADDE	6	6.22
IADDE	7	7.37	ICDEH	6	49.95	IEH	5	48.15
IADDE	6	5.97	IADDE	5	5.66	IEH	4	159.48
IADDE	3	289.61	IADDE	4	27.44	IADDE	5	18.56
IEH	4	5.03	ICDGH	5	57.06	IADDE	6	18.51
IADDE	5	21.64	ICDGH	4	179.41	ICH	3	183.10
IADDE	4	20.21	IADDE	5	16.51	IECH	4	56.42
IADDE	3	54.09	IADDE	4	16.71	IADDE	3	31.97
IADDE	2	454.49			.00			.00

Processamento 08/057 H.M.S

Tabela 9 - Todas as regressões possíveis fornecidas pelo programa MALLOW'S para o cultivar FLINT.

Variáveis	p	Cp	Variáveis	p	Cp	Variáveis	p	Cp
Herituma	1	910,53	IA	2	117,01	IAB	3	90,76
IB	2	164,60	IAC	3	93,95	IABC	4	41,22
IAC	3	40,60	IACD	4	162,54	ICD	3	122,80
IACD	4	10,46	IACDE	5	11,03	IACD	4	59,66
IAD	3	57,17	IACDE	4	9,03	IAD	3	8,56
IAD	2	142,43	IACDE	3	137,86	IAD	4	9,67
IADDE	5	9,76	IACDE	4	52,85	IADDE	5	54,22
IACDE	6	11,73	IACDE	5	11,33	IACDE	4	113,89
ICE	3	154,29	IACDE	4	41,88	IACDE	5	42,33
IACE	4	89,79	IACDE	3	165,94	IACE	4	20,38
IAC	3	111,17	IACDE	2	845,34	IAC	3	314,61
IACF	4	66,51	IACDEF	5	55,86	IACF	4	102,62
IACDF	5	91,77	IACDEF	6	43,71	IACDF	5	42,29
ICDEF	4	141,59	IACDEF	5	109,13	ICDEF	6	13,33
IACDEF	7	13,55	IACDEF	6	54,25	IACDEF	5	52,81
IACDEF	6	11,63	IACDEF	5	11,59	IACDEF	4	134,19
ICF	3	133,46	IACDF	4	10,07	ICF	5	10,34
ICDF	4	53,16	IACDF	5	52,55	ICDF	6	11,72
IACDF	5	12,06	IACDF	4	124,51	IACDF	3	164,40
IACF	4	42,55	IACDF	3	43,17	IACF	4	91,87
ICF	3	105,47	IACDF	4	61,68	ICF	3	80,79
IF	2	330,52	IACDF	3	371,06	IF	4	76,86
IACFG	5	48,56	IACFG	4	91,41	IACFG	5	75,50
IACCFG	6	27,54	IACFG	5	23,49	IACCFG	4	162,93
ICDFG	5	101,57	IACFG	6	12,72	ICDFG	7	13,03
ICDFG	6	52,30	IACFG	5	55,15	ICDFG	6	12,34
IACFG	5	11,65	IACFG	4	103,66	IACFG	5	103,60
IACFG	6	13,10	IACFG	7	13,62	IACFG	6	54,81
ICDFG	7	36,22	IACFG	8	11,63	ICDFG	7	12,88
ICDFG	6	90,49	IACFG	5	117,19	ICDFG	6	16,31
IACDFG	7	17,09	IACFG	6	62,75	IACDFG	5	91,68
IACDFG	6	45,78	IACFG	5	64,78	IACDFG	4	302,99
IEG	3	715,19	IACFG	4	77,57	IEG	5	48,60
IACG	4	98,93	IACG	5	73,58	IACG	6	32,02
IACG	5	35,91	IACG	4	155,69	IACG	5	106,12
IACG	6	13,22	IACG	7	13,70	IACG	6	55,60
IACG	5	54,33	IACG	6	11,73	IACG	5	11,42
IDG	4	118,61	IACG	3	129,21	IDG	4	10,52
IACG	5	10,84	IACG	4	57,26	IACG	5	58,99
IACG	6	12,84	IACG	5	12,44	IACG	4	118,92
ICG	3	162,95	IACG	4	34,00	ICG	5	30,12
IACG	4	74,07	IACG	3	97,04	IACG	4	48,58
IACG	3	84,65	IACG	2	775,28	IACG	3	702,75
IACG	4	36,22	IACG	1	25,76	IACG	4	98,36
IACG	5	74,33	IACG	6	17,81	IACG	5	18,30
ICGH	4	145,07	IACG	5	84,15	ICGH	6	9,37
IACGH	7	10,31	IACG	6	52,21	IACGH	5	50,24
IACGH	6	8,35	IACG	5	7,55	IACGH	4	85,34
IACGH	5	75,64	IACG	6	9,27	IACGH	7	10,64
IACGH	6	46,37	IACG	7	48,10	IACGH	8	11,96
IACGH	7	11,01	IACG	6	73,22	IACGH	5	139,47
IACGH	6	20,31	IACG	7	19,77	IACGH	6	73,89
IACGH	5	100,83	IACG	6	25,71	IACGH	5	33,11
IEGH	4	644,68	IACG	5	292,99	IEGH	6	29,57
IACFGH	7	25,49	IACFGH	6	93,61	IACFGH	7	55,66
IACFGH	8	14,66	IACFGH	7	13,55	IACFGH	6	82,04
ICDFGH	7	62,53	IACFGH	8	9,73	ICDFGH	8	9,00
ICDFGH	8	36,34	IACFGH	7	47,71	ICDFGH	8	11,84
IACDFGH	7	10,85	IACFGH	6	72,36	IACDFGH	5	74,62
IACDFGH	6	8,91	IACFGH	7	9,96	IACDFGH	6	49,20
ICDFGH	7	45,53	IACFGH	8	11,52	ICDFGH	7	10,65
ICDFGH	6	70,85	IACFGH	5	141,35	ICDFGH	6	17,18
IACFGH	7	17,37	IACFGH	6	75,09	IACFGH	5	93,38
IACFGH	6	26,81	IACFGH	5	35,15	IACFGH	4	359,76
IFH	3	373,82	IACFGH	4	56,27	IFH	5	51,73
IFH	4	106,07	IACFGH	5	87,26	IFH	6	40,90
IACFH	5	39,35	IACFH	4	140,61	IACFH	5	103,83
IACFH	6	8,66	IACFH	7	9,60	IACFH	6	49,49
ICDFH	5	48,42	IACFH	6	9,17	ICDFH	5	7,54
ICDFH	4	116,77	IACFH	5	116,36	ICDFH	6	9,38
IACDFH	7	10,91	IACDFH	6	47,27	IACDFH	7	48,87
IACDFH	8	10,97	IACDFH	7	10,37	IACDFH	6	76,79
ICFFH	5	103,71	IACDFH	6	41,27	ICFFH	7	42,88
IACFFH	6	81,90	IACDFH	5	103,15	IACFFH	6	44,94
IACFFH	5	45,49	IACDFH	4	309,23	IACFFH	3	693,57
IACFFH	4	94,72	IACDFH	5	85,99	IACFFH	4	155,90
ICEH	5	83,05	IACDFH	6	41,04	ICEH	5	59,32
ICEH	4	137,82	IACDFH	5	37,15	ICEH	6	16,13
IACDEH	7	11,65	IACDEH	6	47,01	IACDEH	5	45,43
IACDEH	6	9,67	IACDEH	5	8,23	IACDEH	4	114,77
IDH	3	117,20	IACDEH	4	7,16	IDH	5	8,79
IDH	4	45,39	IACDEH	5	50,33	IDH	6	10,79
IACDH	5	9,14	IACDH	4	103,03	IACDH	3	143,13
IACDH	4	37,88	IACDH	5	39,77	IACDH	4	85,82
IDH	3	154,19	IACDH	4	86,43	IDH	3	98,82
IH	2	753,82	IACDH	1	0,01	IH	1	0,60

Processamento 08/2024 H.N.S

Tabela 10 - Todas as regressões possíveis fornecidas pelo programa MALLOW'S para o cultivar PIRANÃO.

Variáveis	n	Cp	Variáveis	n	Cp	Variáveis	n	Cp
Nenhuma	1	633,63	IA	2	56,56	IAE	3	38,07
IA	2	99,07	IAE	3	71,04	IAEC	4	18,55
IAE	3	24,96	IAEC	4	271,07	IAECD	5	117,05
IAEC	4	6,31	IAECD+	5	3,32	IAECDE	6	35,48
IAECD+	5	33,63	IAEDE+	6	2,18	IAE	3	6,00
IAEDE+	6	117,61	IAE	3	117,03	IAED	4	7,99
IAED	4	4,14	IAED	4	35,54	IAEDE	5	37,26
IAEDE	5	5,31	IAEDE	5	8,30	IAEDE+	6	115,89
IAEDE+	6	293,86	IAE	3	22,29	IAE	3	16,47
IAE	3	65,96	IAE	3	76,78	IAE	3	24,03
IAE	3	36,39	IAE	3	419,45	IAE	3	315,74
IAE	3	28,20	IAE	3	18,87	IAE	3	67,38
IAE	3	67,93	IAE	3	17,55	IAE	3	20,42
IAE	3	211,49	IAE	3	113,04	IAE	3	9,49
IAE	3	7,30	IAE	3	38,47	IAE	3	37,05
IAE	3	5,44	IAE	3	9,14	IAE	3	118,17
IAE	3	113,32	IAE	3	7,19	IAE	3	3,54
IAE	3	35,10	IAE	3	36,56	IAE	3	5,38
IAE	3	7,79	IAE	3	118,38	IAE	3	211,19
IAE	3	18,49	IAE	3	16,24	IAE	3	69,14
IAE	3	83,22	IAE	3	29,01	IAE	3	42,13
IAE	3	433,26	IAE	3	334,88	IAE	3	31,46
IAE	3	15,67	IAE	3	51,98	IAE	3	48,13
IAE	3	9,84	IAE	3	14,66	IAE	3	183,07
IAE	3	115,63	IAE	3	9,79	IAE	3	6,55
IAE	3	37,82	IAE	3	36,01	IAE	3	5,35
IAE	3	9,01	IAE	3	113,89	IAE	3	115,84
IAE	3	10,78	IAE	3	7,34	IAE	3	37,34
IAE	3	39,75	IAE	3	7,68	IAE	3	18,18
IAE	3	114,66	IAE	3	131,59	IAE	3	9,09
IAE	3	6,32	IAE	3	39,84	IAE	3	53,94
IAE	3	16,68	IAE	3	28,95	IAE	3	388,05
IAE	3	415,66	IAE	3	37,81	IAE	3	22,61
IAE	3	64,49	IAE	3	46,84	IAE	3	11,53
IAE	3	20,32	IAE	3	228,51	IAE	3	115,51
IAE	3	18,28	IAE	3	6,98	IAE	3	37,87
IAE	3	36,39	IAE	3	6,14	IAE	3	9,31
IAE	3	112,96	IAE	3	112,02	IAE	3	7,56
IAE	3	4,16	IAE	3	35,02	IAE	3	36,47
IAE	3	5,68	IAE	3	8,28	IAE	3	113,73
IAE	3	226,95	IAE	3	18,33	IAE	3	9,75
IAE	3	46,77	IAE	3	63,09	IAE	3	22,87
IAE	3	44,22	IAE	3	497,25	IAE	3	228,37
IAE	3	37,25	IAE	3	14,30	IAE	3	38,98
IAE	3	31,98	IAE	3	10,74	IAE	3	28,32
IAE	3	184,13	IAE	3	94,68	IAE	3	10,27
IAE	3	6,19	IAE	3	24,49	IAE	3	23,32
IAE	3	4,42	IAE	3	-9,15	IAE	3	94,70
IAE	3	96,48	IAE	3	10,95	IAE	3	6,42
IAE	3	24,37	IAE	3	26,06	IAE	3	8,13
IAE	3	12,26	IAE	3	96,58	IAE	3	186,08
IAE	3	22,31	IAE	3	12,63	IAE	3	33,27
IAE	3	32,84	IAE	3	15,81	IAE	3	34,02
IAE	3	211,16	IAE	3	210,17	IAE	3	30,46
IAE	3	15,98	IAE	3	34,78	IAE	3	27,32
IAE	3	7,64	IAE	3	11,08	IAE	3	188,57
IAE	3	95,82	IAE	3	12,16	IAE	3	9,06
IAE	3	27,71	IAE	3	25,92	IAE	3	8,32
IAE	3	12,69	IAE	3	95,63	IAE	3	93,86
IAE	3	18,20	IAE	3	6,33	IAE	3	24,87
IAE	3	26,39	IAE	3	6,12	IAE	3	11,75
IAE	3	95,85	IAE	3	148,00	IAE	3	16,61
IAE	3	10,76	IAE	3	32,73	IAE	3	32,77
IAE	3	14,42	IAE	3	32,53	IAE	3	218,10
IAE	3	225,33	IAE	3	32,18	IAE	3	13,08
IAE	3	30,93	IAE	3	30,88	IAE	3	9,65
IAE	3	16,26	IAE	3	150,59	IAE	3	115,35
IAE	3	9,76	IAE	3	6,15	IAE	3	27,08
IAE	3	25,14	IAE	3	4,45	IAE	3	9,19
IAE	3	116,44	IAE	3	113,10	IAE	3	11,11
IAE	3	6,43	IAE	3	25,00	IAE	3	25,74
IAE	3	7,61	IAE	3	11,11	IAE	3	95,15
IAE	3	114,13	IAE	3	16,56	IAE	3	10,42
IAE	3	28,05	IAE	3	32,75	IAE	3	14,20
IAE	3	23,57	IAE	3	224,68	IAE	3	238,52
IAE	3	32,64	IAE	3	13,82	IAE	3	31,02
IAE	3	31,37	IAE	3	12,02	IAE	3	22,41
IAE	3	185,42	IAE	3	110,66	IAE	3	10,38
IAE	3	6,13	IAE	3	25,04	IAE	3	23,05
IAE	3	4,69	IAE	3	9,90	IAE	3	111,17
IAE	3	115,02	IAE	3	7,99	IAE	3	2,76
IAE	3	23,15	IAE	3	25,15	IAE	3	4,23
IAE	3	8,30	IAE	3	115,66	IAE	3	183,74
IAE	3	21,74	IAE	3	10,18	IAE	3	29,94
IAE	3	29,25	IAE	3	12,35	IAE	3	35,28
IAE	3	238,08	IAE	3	.00	IAE	3	.00

Tabela 11 - Vinte primeiras equações em ordem crescente de C_p 's.

ORDEM	C-111		FLINT		PIRANÃO	
	Equação	C_p	Equação	C_p	Equação	C_p
1	ACE	3,67*	ADH	7,16	ABD	2,18*
2	ACEF	4,02*	ADFH	7,54	ABDH	2,76*
3	ACDE	4,65*	ADGH	7,55	ABCD	3,32*
4	ACEFH	4,78*	ADEH	8,23	ABDF	3,54*
5	ABEH	4,79*	ABDGH	8,35	ABDE	4,14*
6	ADE	4,84	AD	8,56	ABDG	4,16*
7	ABEGH	5,08*	ACDFH	8,66	ABCDH	4,23*
8	ACEH	5,30	ABDH	8,79	ABDGH	4,42*
9	ACEG	5,33	ADFGH	8,91	ABDFH	4,45*
10	ABCE	5,42	ABCDEFGH	9,00*	ABDEH	4,69*
11	ADEH	5,66	ABD	9,03	ABCDG	5,00*
12	ADEF	5,73	ACDH	9,14	ABCDF	5,30*
13	ADEGH	5,74*	ABDFH	9,17	ABCDE	5,31*
14	ABCEH	5,78*	ADEGH	9,27	ABDFG	5,35*
15	ACDFG	5,79*	ACDGH	9,37	ABDEF	5,44*
16	ADEG	5,93	ADEFH	9,38	AD	6,00
17	ABDEH	5,97*	ABCDFH	9,60	ABCDEH	6,13*
18	ABCEF	5,98*	ADE	9,67	ABDEG	6,14
19	ACEFG	5,98*	ABDEH	9,67	ABCDFH	6,15*
20	ACDEF	6,01	ACDEFGH	9,73	ABCDGH	6,19*

* Equações com $C_p \leq p$.

Os valores dos C_p 's calculados para o cultivar FLINT são notadamente maiores que os C_p 's calculados para os demais cultivares, o que certamente deve-se ao fato de a equação completa ter apresentado um maior número de parâmetros significativos, forçando para cima os valores de C_p 's.

Estas constatações importantes demonstram efetivamente que o método de seleção de todas as possíveis regressões oferece bem mais informações sobre as variáveis do que os métodos "stepwise".

O tempo gasto no processamento das 256 equações foi de aproximadamente 20 minutos. Levando-se em conta a não compilação do programa e a baixa velocidade de impressão do equipamento, podemos considerar viável a utilização do microcomputador PET/CBM neste tipo de análise, desde que o número de variáveis independentes não seja maior do que 10.

As equações que apresentaram menores valores de C_p 's em cada cultivar foram analisadas para obtenção dos parâmetros e estudo dos resíduos. Os resultados obtidos são apresentados nas Tabelas 12, 13 e 14.

Mesmo com menos variáveis, as equações mantiveram elevadas as porcentagens de variação explicada (em torno de 97%), o que fez com que os valores de F das análises de variância aumentassem significativamente. Para os pesquisadores mais incautos, estes resultados altamente significativos já seriam suficientes para que as equações fossem dadas como satisfatórias. A análise dos resíduos

nestas circunstâncias apresentam relevantes informações com respeito aos modelos propostos. Observa-se que nenhum ponto discrepante ("outlier") pode ser constatado nas colunas referentes aos resíduos padronizados e ordenados em relação a \bar{Y}_j . Isto reforça a idéia de que, se alguma tendência ocorrer nos resíduos, será certamente proveniente de uma formulação errônea do modelo, e não da condição de dados inadequados ou mal coletados. O pequeno número de variáveis nas equações pode ser devido à presença da variável NINS, representativa da população de insetos, e que sem dúvida influencia nas demais variáveis, podendo, talvez, justificar a elevada multicolinearidade observada. As variações sofridas pelas variáveis químicas, por sua vez, modificam o comportamento da população de insetos no decorrer do tempo, retroalimentando o sistema.

Tabela 12 - Resultados do programa STEP.COMP, para análise de regressão da equação ACE, selecionada para o cultivar C.111.

Desvio Padrão Residual= 1.75517762
 Coeficiente de Correlação= .982980705
 % de Variação Explicada= 96.6251067
 Termo Constante= 77.3654388

Análise da Variância

C.Variacao	G.L.	S.Q.	Q.M.	F
Resressao	3	2822.4225	940.80751	305.39
Residuo	32	98.5807	3.08064	
Total	35	2921.0033		

Variavel	Coeficiente B	Erro Padrao	Intervalo de Confianca
NHS	.978648 *	.07662	.82264 <=B 1<= 1.13465
CARB	1.062963- *	.44907	1.97727- <=B 3<= .14865-
MASE	3.438760 *	.82354	1.76202 <=B 5<= 5.11549

VALORES ORDENADOS EM RELACAO A Y.est.

Obs	Y.obs.	Y.est.	Residuo/s.	Obs	Residuo
2	5.4376	3.8765	+ .88941	2	+ 1.56107
3	5.4255	4.1922	+ .70261	3	+ 1.23321
25	4.0098	4.6953	- .39052	25	- .68544
27	5.2397	4.7491	+ .27952	27	+ .49062
13	6.8122	4.7754	+ 1.16048	13	+ 2.03685
14	6.6324	4.8146	+ 1.03566	14	+ 1.81777
1	4.3601	4.8298	- .26756	1	- .46962
26	4.8878	5.2271	- .19328	26	- .33925
15	7.9270	5.3461	+ 1.47050	15	+ 2.58099
6	7.6497	9.0695	- .80891	6	- 1.41978
5	8.1484	10.4719	- 1.32376	5	- 2.32344
18	9.8423	11.6931	- 1.05447	18	- 1.85079
4	9.4185	12.1397	- 1.55034	4	- 2.72112
8	12.4856	13.6611	- .66973	8	- 1.17550
16	13.6448	14.5017	- .48820	16	- .85689
11	14.5860	15.7244	- .64858	11	- 1.13838
17	17.4846	16.7450	+ .42135	17	+ .73954
12	16.1189	17.1432	- .58358	12	- 1.02428
9	14.9406	17.1518	- 1.25980	9	- 2.21118
7	15.3708	17.9119	- 1.44773	7	- 2.54103
10	17.2765	17.9852	- .40375	10	- .70866
30	22.7471	19.8063	+ 1.67551	30	+ 2.94082
28	23.0850	20.2453	+ 1.61788	28	+ 2.83967
22	22.6401	21.9010	+ .42104	22	+ .73901
19	22.9911	22.5326	+ .23272	19	+ .40847
29	24.8067	23.1078	+ .96795	29	+ 1.69893
35	24.6461	24.5795	+ .03793	35	+ .06658
31	22.3173	24.6059	- 1.30390	31	- 2.28858
20	25.8141	25.7234	+ .05168	20	+ .09071
21	25.7183	25.9200	- .11491	21	- .20169
32	24.0881	26.1677	- 1.18494	32	- 2.07961
33	30.1479	26.8338	+ 1.88822	33	+ 3.31417
23	31.5022	30.0272	+ .84040	23	+ 1.47506
34	31.0963	30.1745	+ .52519	34	+ .92181
24	30.3612	30.7928	- .24592	24	- .43163
36	30.4327	30.9212	- .27827	36	- .48841

Tabela 13 - Resultados do programa STEP.COMP para análise de regressão da equação ADH, selecionada para o cultivar FLINT.

Desvio Padrão Residual= 1.38844652
 Coeficiente de Correlação= .981212262
 % de Variação Explicada= 96.2777503
 Termo Constante=-49.0942975

Análise da

C. Variação	G.L.	S.Q.	Q.M.	F
Regressão	3	1595.6172	531.87241	275.89
Resíduo	32	61.6890	1.92778	
Total	35	1657.3063		

Variável	Coefficiente B	Erro Padrão	Intervalo de Confiança
NINS	.717715 *	.07107	.57300 <=B 1<= .86242
UMID	4.380832 *	.47448	3.41478 <=B 4<= 5.34688
IACD	.026417-	.01500	.05697-<=B 8<= .00413

VALORES ORDENADOS EM RELAÇÃO A Y.est.

Obs	Y.obs.	Y.est.	Resíduo/s	Obs	Resíduo
1	4.3146	2.3859	+ 1.38904	1	+ 1.92861
3	4.5921	3.3999	+ .85865	3	+ 1.19220
2	5.3673	3.5568	+ 1.30399	2	+ 1.81053
15	7.6366	6.1823	+ 1.04744	15	+ 1.45432
14	7.3214	6.4513	+ .62670	14	+ .87014
4	6.3986	6.5395	- .10153	4	- .14097
25	6.4475	7.5664	- .80580	25	- 1.11881
13	8.1770	7.6459	+ .38247	13	+ .53105
26	7.1564	8.0640	- .65368	26	- .90760
27	7.6061	8.6904	- .78096	27	- 1.08432
8	10.2165	10.0840	+ .09542	8	+ .13249
5	10.8618	10.4932	+ .26548	5	+ .36861
6	9.3901	11.0254	- 1.17779	6	- 1.63531
18	10.4888	11.4455	- .68903	18	- .95668
20	11.9226	11.9955	- .05246	20	- .07284
12	13.1568	12.5800	+ .41546	12	+ .57685
9	10.5048	12.8395	- 1.68153	9	- 2.33471
11	12.8286	13.7113	- .63575	11	- .88271
7	11.6837	14.0591	- 1.71084	7	- 2.37541
17	15.1508	15.7800	- .45313	17	- .62914
19	15.8827	15.9127	+ .69858	19	+ .96994
16	16.6743	16.0918	+ .41955	16	+ .58253
23	15.8080	16.1234	- .22717	23	- .31542
21	16.3420	16.1906	+ .10903	21	+ .15138
10	14.9980	16.6592	- 1.19649	10	- 1.66126
24	16.6472	17.4150	- .55300	24	- .76781
29	16.0414	18.1356	- 1.50826	29	- 2.09414
22	17.5325	18.5619	- .74138	22	- 1.02937
32	19.3825	19.2333	+ .10746	32	+ .14921
31	23.7799	21.0914	+ 1.93627	31	+ 2.68841
35	22.8554	21.2814	+ 1.13362	35	+ 1.57398
30	21.3936	21.5942	- .14448	30	- .20061
33	26.1154	23.2768	+ 2.04444	33	+ 2.83859
28	22.9041	24.0355	- .81484	28	- 1.13136
34	24.6756	24.0544	+ .44736	34	+ .62113
36	32.6918	31.7933	+ .64713	36	+ .89851

Tabela 14 - Resultados do programa STEP.COMP para análise de regressão da equação ABD, selecionada para o cultivar PIRANÃO.

Desvio Padrão Residual= 1.56352911
 Coeficiente de Correlação= .976888663
 % de Variação Explicada= 95.8223014
 Termo Constante=-44.4837788

Análise

C.Variação	G.L.	S.Q.	Q.M.	F
Regressão	3	1794.2843	598.09497	244.65
Resíduo	32	78.2279	2.44462	
Total	35	1872.5128		

Variável	Coeficiente B	Erro Padrão	Intervalo de Confiança
NINS	.538095 *	.09037	.35408 <=B 1<= .72210
TARM	.025878 *	.01042	.00465 <=B 2<= .04709
UMID	3.940090 *	.62173	2.67411 <=B 4<= 5.20606

VALORES ORDENADOS EM RELAÇÃO A Y.est.

Obs	Y.obs.	Y.est.	Resíduo/s	Obs	Resíduo
1	4.7613	3.5521	+ .77342	1	+ 1.20926
3	4.3601	3.6190	+ .47399	3	+ .74110
2	4.1312	3.6663	+ .29733	2	+ .46489
27	5.7247	5.0611	+ .42442	27	+ .66359
15	7.1587	5.7309	+ .91319	15	+ 1.42780
26	6.0331	5.9870	+ .06144	26	+ .09607
13	6.9710	5.9988	+ .62174	13	+ .97211
14	6.8025	6.0186	+ .50140	14	+ .78395
25	5.8021	6.2510	- .28710	25	- .44889
9	8.7306	8.7823	- .03304	9	- .05167
4	10.1672	10.2170	- .03188	4	- .04985
6	10.1209	11.5203	- .89501	6	- 1.39938
11	11.2694	11.9074	- .40809	11	- .63806
8	11.7629	13.0343	- .81319	8	- 1.27146
7	11.3943	13.4321	- 1.30337	7	- 2.03786
12	11.9552	13.5084	- .99338	12	- 1.55319
10	11.1552	13.5827	- 1.55256	10	- 2.42747
18	17.2998	14.4945	+ 1.79418	18	+ 2.80526
5	12.7623	14.6617	- 1.21481	5	- 1.89939
16	14.4857	14.6957	- .13430	16	- .20998
21	12.6448	15.1683	- 1.61403	21	- 2.52359
17	15.4357	16.7754	- .85684	17	- 1.33970
20	17.2482	16.9161	+ .21243	20	+ .33215
28	18.0066	17.6878	+ .20393	28	+ .31885
33	17.0224	18.0449	- .65398	33	- 1.02252
30	17.1336	18.2827	- .73497	30	- 1.14915
22	19.9859	18.5605	+ .91163	22	+ 1.42537
19	20.4181	18.6658	+ 1.12078	19	+ 1.75237
24	21.5643	18.7530	+ 1.79806	24	+ 2.81133
23	17.4275	19.5237	- 1.34071	23	- 2.09624
31	21.9397	20.0677	+ 1.19725	31	+ 1.87194
32	23.0015	22.5501	+ .28869	32	+ .45138
29	23.4897	23.3785	+ .07110	29	+ .11117
35	29.2714	26.7385	+ 1.61995	35	+ 2.53285
36	26.6601	28.6136	- 1.24941	36	- 1.95349
34	32.1639	30.8635	+ .83172	34	+ 1.30042

Os resíduos resultantes destas equações foram estudados graficamente contra os valores de \bar{Y}_i e contra a variável NINS, considerada de grande relevância. Os gráficos são apresentados na Figura 4.

Os gráficos 4.a, 4.b e 4.c, relativos aos resíduos dos cultivares contra \bar{Y}_i apresentam uma tendência acentuadamente semelhante à Figura 2.h, indicativa de não linearidade e variância não constante dos resíduos.

A mesma tendência, embora menos acentuada, é verificada nos gráficos 4.d, 4.e e 4.f, para os resíduos estudados contra a variável NINS. Uma explicação bem plausível é que, sendo a variável NINS aquela que mais influencia nas perdas de peso, o seu comportamento reconhecidamente não linear transferiu-se para o modelo.

Embora não constados aqui, várias outras equações entre aquelas 20 consideradas foram também analisadas, e todas elas apresentaram comportamento para os resíduos iguais aos aqui relatados. Certamente isto ocorreu pela constante presença da variável NINS em todas as equações.

Um recurso seria a escolha da melhor equação entre as 256 avaliadas para cada cultivar, que não contenha a variável NINS, embora sabendo de antemão que a porcentagem de variação explicada será drasticamente reduzida, por se eliminar aquela mais explicativa. No entanto, a atitude mais coerente seria a reformulação do modelo, não mais utilizando aquele de regressão linear múltipla, vi

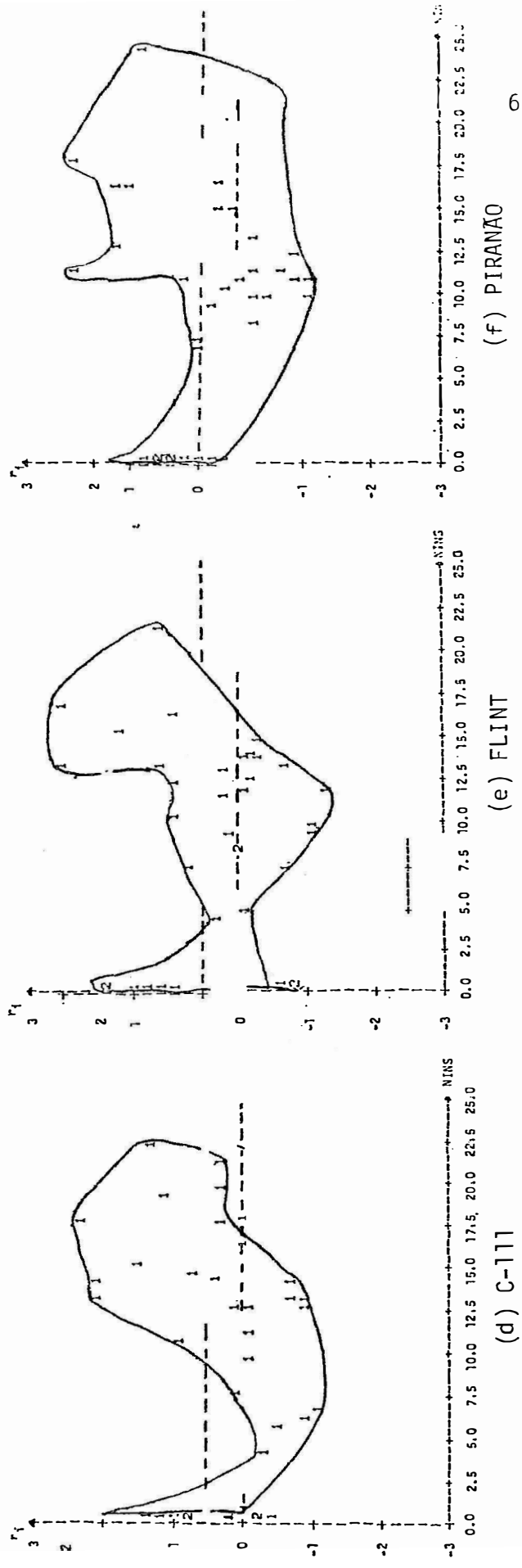
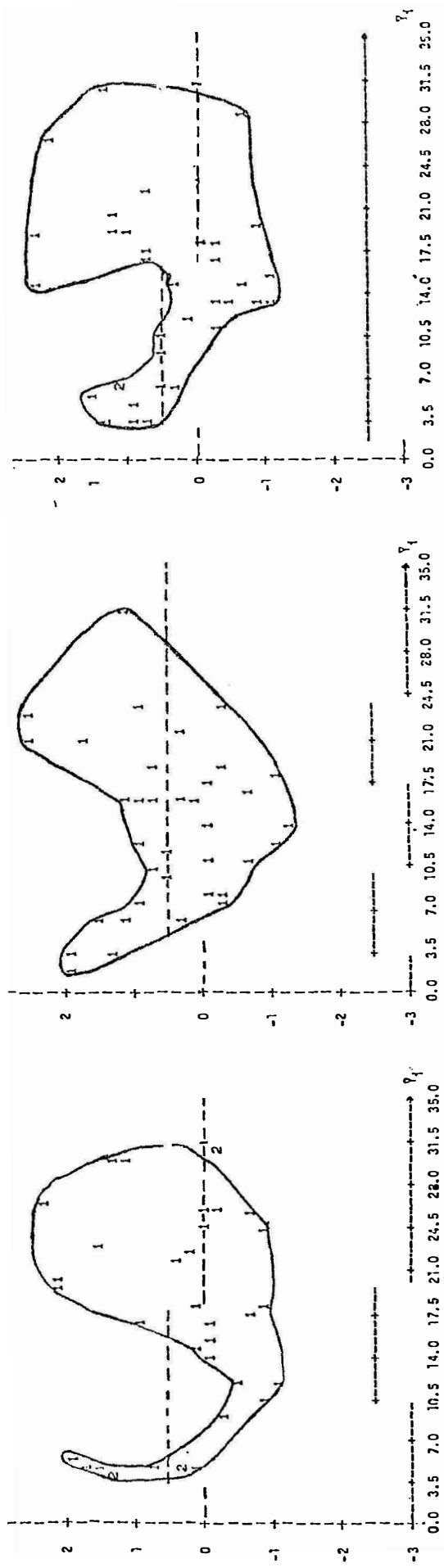


Figura 4 - Gráficos dos resíduos contra \hat{Y}_i e NINS, para os três cultivares de milho, fornecidos pelo programa PLOT.RES.

sualmente inadequado e, sim, algum outro que avalie verdadeiramente o comportamento da população de insetos. Normalmente, para este tipo de dados, modelos de simulação apresentam resultados bem mais realistas.

É também razoável esperar que resolvido o problema de não linearidade, a heterocedasticidade também deixa de existir.

Embora apenas as constatações visuais dos diagramas de dispersão dos resíduos sejam suficientes para indicar anomalias nas equações, o que confirma as opiniões de WEISBERG (1980) e DRAPER e SMITH (1966), obtivemos também os resultados para os testes dos chorrilhos e Lilliefors.

Para o primeiro teste foram observadas sequências de sinais dos resíduos ordenados em relação a \hat{Y}_i e apresentados nas Tabelas 12, 13 e 14.

Estas sequências totalizam $r=14$ chorrilhos para o C-111, $r=17$ para o FLINT e $r=11$ para o PIRANÃO. As distribuições nulas para os n_1 's sinais negativos e n_2 's positivos são apresentadas na Tabela 15, onde os números de chorrilhos obtidos são assinalados por uma seta.

Para o C-111 observa-se $P(r \leq 14) < 0,066$, inferior portanto ao nível $\alpha = 0,1$ estabelecido, levando-nos a considerar os erros autocorrelacionados. O cultivar FLINT apresentou um resultado satisfatório para este teste, pois $P(r \leq 17) < 0,31$, bem superior ao nível α , portanto, para este cultivar, consideramos os erros não autocorrelacionados.

Tabela 15 - Distribuições nulas dos resíduos para os três cultivares de milho, geradas pelo programa STEP.COMP.

C-111

r=14 , N₁=19 , N₂=17

r	P(r<=r.)	P(r<=r.)	F(r>=r.)
02	0.00000	0.00000	1.00000
03	0.00000	0.00000	1.00000
04	0.00000	0.00000	0.99999
05	0.00000	0.00000	0.99999
06	0.00000	0.00000	0.99999
07	0.00002	0.00002	0.99997
08	0.00010	0.00010	0.99986
09	0.00037	0.00037	0.99949
10	0.00129	0.00129	0.99819
11	0.00326	0.00326	0.99483
12	0.00870	0.00870	0.98612
13	0.01741	0.01741	0.96871
14	0.03459	0.03459	0.93413
15	0.05434	0.05434	0.87978
16	0.06459	0.06459	0.79599
17	0.10586	0.10586	0.68523
18	0.13100	0.13100	0.55522
19	0.13938	0.13938	0.42721
20	0.10351	0.10351	0.29732
21	0.08151	0.08151	0.19431
22	0.05167	0.05167	0.11250
23	0.03233	0.03233	0.06092
24	0.01616	0.01616	0.02859
25	0.00785	0.00785	0.01242
26	0.00302	0.00302	0.00456
27	0.00111	0.00111	0.00154
28	0.00031	0.00031	0.00042
29	0.00006	0.00006	0.00010
30	0.00000	0.00000	0.00002
31	0.00000	0.00000	0.00000
32	0.00000	0.00000	0.00000
33	0.00000	0.00000	0.00000
34	0.00000	0.00000	0.00000
35	0.00000	0.00000	0.00000
36	0.00000	1.00000	0.00000

FLINT

r=17 , N₁=18 , N₂=18

r	P(r<=r.)	P(r<=r.)	P(r>=r.)
02	0.00000	0.00000	1.00000
03	0.00000	0.00000	1.00000
04	0.00000	0.00000	0.99999
05	0.00000	0.00000	0.99999
06	0.00000	0.00000	0.99999
07	0.00002	0.00002	0.99997
08	0.00010	0.00010	0.99986
09	0.00035	0.00035	0.99951
10	0.00124	0.00124	0.99916
11	0.00324	0.00324	0.99626
12	0.00843	0.00843	0.98562
13	0.01687	0.01687	0.96558
14	0.03375	0.03375	0.92570
15	0.05364	0.05364	0.83595
16	0.08335	0.08335	0.68250
17	0.10419	0.10419	0.52255
18	0.13624	0.13624	0.36536
19	0.13024	0.13024	0.25512
20	0.10419	0.10419	0.16487
21	0.08335	0.08335	0.10453
22	0.06530	0.06530	0.06044
23	0.05304	0.05304	0.11709
24	0.03375	0.03375	0.05404
25	0.01687	0.01687	0.03029
26	0.00843	0.00843	0.01341
27	0.00324	0.00324	0.00457
28	0.00124	0.00124	0.00173
29	0.00035	0.00035	0.00048
30	0.00010	0.00010	0.00012
31	0.00002	0.00002	0.00000
32	0.00000	0.00000	0.00000
33	0.00000	0.00000	0.00000
34	0.00000	0.00000	0.99999
35	0.00000	0.99999	0.00000
36	0.00000	1.00000	0.00000

PIRANÃO

r=11 , N₁=17 , N₂=19

r	P(r<=r.)	P(r<=r.)	P(r>=r.)
02	0.00000	0.00000	1.00000
03	0.00000	0.00000	1.00000
04	0.00000	0.00000	0.99999
05	0.00000	0.00000	0.99999
06	0.00000	0.00000	0.99999
07	0.00002	0.00002	0.99997
08	0.00010	0.00010	0.99986
09	0.00037	0.00037	0.99956
10	0.00129	0.00129	0.99918
11	0.00326	0.00326	0.99813
12	0.00870	0.00870	0.99483
13	0.01741	0.01741	0.98512
14	0.03459	0.03459	0.96587
15	0.05434	0.05434	0.93413
16	0.08459	0.08459	0.87978
17	0.10586	0.10586	0.79599
18	0.13100	0.13100	0.68523
19	0.13938	0.13938	0.55522
20	0.10351	0.10351	0.42721
21	0.08151	0.08151	0.29732
22	0.05167	0.05167	0.19431
23	0.03233	0.03233	0.11250
24	0.01616	0.01616	0.06092
25	0.00785	0.00785	0.02859
26	0.00302	0.00302	0.01242
27	0.00111	0.00111	0.00456
28	0.00031	0.00031	0.00154
29	0.00006	0.00006	0.00042
30	0.00000	0.00000	0.00010
31	0.00000	0.00000	0.00002
32	0.00000	0.00000	0.00000
33	0.00000	0.00000	0.00000
34	0.00000	0.00000	0.99999
35	0.00000	1.00000	0.00000
36	0.00000	1.00000	0.00000

Para o PIRANÃO foi considerada acentuada a autocorrelação dos erros, pois $P(r \leq 11) \leq 0,0052$.

Estes resultados se aproximam das conclusões extraídas das análises gráficas, excetuando-se para o cultivar FLINT, que não acusou tendências na distribuição dos resíduos pelo presente teste.

Os resultados obtidos para o teste de Lilliefors, apresentado nas Tabelas 16, 17 e 18 foram satisfatórios, levando-nos a considerar, para os três cultivares, que os erros são normalmente distribuídos. Para $N=36$ e $\alpha = 0,05$, o valor d de tabela é $\frac{0,886}{\sqrt{36}} = 0,1477$, e os D 's na ordem, C-111, FLINT e PIRANÃO foram 0,0766, 0,0767, 0,0709, respectivamente, notadamente similares. Constata-se, nestas tabelas, que os $r_i^!$ estão bem dentro da amplitude -2 a $+2$, concordando com as afirmações de DRAPER e SMITH (1966).

O fato de ter sido constatada acentuada multicolinearidade nos fez supor ser ela, talvez, uma das fontes dos distúrbios nas equações ajustadas. Baseados nesta suposição, procedemos à análise dos dados, utilizando o método dos componentes principais sugerido por CHATTERJEE e PRICE (1977), para estas situações.

As análises foram realizadas em duas etapas, sendo a primeira pelo programa AUTO-JACOBI, o qual processa os seguintes cálculos:

- 1) Padroniza as variáveis independentes.
- 2) Obtém as matrizes de correlações para estas variáveis.

Tabela 16 - Resultados do teste de Lilliefors para normalidade dos resíduos do cultivar C-111, fornecidos pelo programa STEP.COMP.

Residuo	r_i	$F(r_i)$	$S(r_i)$	$ F(r_i)-S(r_i) $	
-	2.72112	-1.550	0.06052	0.02777	0.06052
-	2.54103	-1.447	0.07384	0.05555	0.04606
-	2.32344	-1.323	0.09279	0.08333	0.03723
-	2.28858	-1.303	0.09613	0.11111	0.01497
-	2.21118	-1.259	0.10386	0.13888	0.03501
-	2.07961	-1.184	0.11803	0.16666	0.04862
-	1.85079	-1.054	0.14583	0.19444	0.04861
-	1.41978	-.808	0.20928	0.22222	0.01483
-	1.17550	-.669	0.25151	0.25000	0.02929
-	1.13838	-.648	0.25830	0.27777	0.01947
-	1.02428	-.583	0.27975	0.30555	0.02580
-	.85689	-.488	0.31270	0.33333	0.02063
-	.70866	-.403	0.34319	0.36111	0.01791
-	.68544	-.390	0.34807	0.38888	0.04081
-	.48841	-.278	0.39040	0.41666	0.02626
-	.46962	-.267	0.39451	0.44444	0.04992
-	.43163	-.245	0.40287	0.47222	0.06935
-	.33925	-.193	0.42336	0.50000	0.07663 ←
-	.20169	-.114	0.45425	0.52777	0.07352
+	.06658	+.037	0.51513	0.55555	0.04042
+	.09071	+.051	0.52061	0.58333	0.06272
+	.40847	+.232	0.59201	0.61111	0.01909
+	.49062	+.279	0.61008	0.63888	0.02880
+	.73901	+.421	0.66313	0.66666	0.02425
+	.73954	+.421	0.66325	0.69444	0.03119
+	.92181	+.525	0.70027	0.72222	0.02194
+	1.23321	+.702	0.75885	0.75000	0.03663
+	1.47506	+.840	0.79965	0.77777	0.04965
+	1.56107	+.889	0.81310	0.80555	0.03533
+	1.69893	+.967	0.83346	0.83333	0.02791
+	1.81777	+1.035	0.84982	0.86111	0.01648
+	2.03685	+1.160	0.87707	0.88888	0.01596
+	2.58099	+1.470	0.92928	0.91666	0.04039
+	2.83967	+1.617	0.94715	0.94444	0.03048
+	2.94882	+1.675	0.95308	0.97222	0.01913
+	3.31417	+1.888	0.97050	0.99999	0.02949

D = 0,0766

Tabela 17 - Resultados do teste de Lilliefors para normalidade dos resíduos do cultivar FLINT, fornecidos pelo programa STEP.COMP.

	Residuo	r_i	$F(r_i)$	$S(r_i)$	$ F(r_i)-S(r_i) $
-	2.52359	-1.614	0.05325	0.02777	0.05325
-	2.42747	-1.552	0.06026	0.05555	0.03248
-	2.09624	-1.340	0.09000	0.08333	0.03445
-	2.03786	-1.303	0.09622	0.11111	0.01488
-	1.95349	-1.249	0.10575	0.13888	0.03313
-	1.89939	-1.214	0.11221	0.16666	0.05444
-	1.55319	-.993	0.16026	0.19444	0.03418
-	1.39938	-.895	0.18538	0.22222	0.03683
-	1.33970	-.856	0.19576	0.25000	0.05423
-	1.27146	-.813	0.20805	0.27777	0.06972
-	1.14915	-.734	0.23117	0.30555	0.07437
-	1.02252	-.653	0.25655	0.33333	0.07677 ←
-	.63806	-.408	0.34160	0.36111	0.01950
-	.44889	-.287	0.38701	0.38888	0.02590
-	.20998	-.134	0.44658	0.41666	0.05769
-	.05167	-.033	0.48681	0.44444	0.07015
-	.04985	-.031	0.48728	0.47222	0.04283
+	.09607	+.061	0.52449	0.50000	0.05227
+	.11117	+.071	0.52834	0.52777	0.02834
+	.31885	+.203	0.58079	0.55555	0.05302
+	.33215	+.212	0.58411	0.58333	0.02856
+	.45138	+.288	0.61359	0.61111	0.03026
+	.46489	+.297	0.61689	0.63888	0.02199
+	.66359	+.424	0.66437	0.66666	0.02548
+	.74110	+.473	0.68224	0.69444	0.01558
+	.78395	+.501	0.69195	0.72222	0.03026
+	.97211	+.621	0.73294	0.75000	0.01705
+	1.20926	+.773	0.78036	0.77777	0.03036
+	1.30042	+.831	0.79721	0.80555	0.01944
+	1.42537	+.911	0.81902	0.83333	0.01431
+	1.42780	+.913	0.81942	0.86111	0.04168
+	1.75237	+1.120	0.86880	0.88888	0.02007
+	1.87194	+1.197	0.88439	0.91666	0.03226
+	2.53285	+1.619	0.94737	0.94444	0.03071
+	2.80526	+1.794	0.96360	0.97222	0.01916
+	2.81133	+1.798	0.96391	0.99999	0.03608

D = 0,0767.

Tabela 18 - Resultados do teste de Lilliefors para normalidade dos resíduos do cultivar PIRANÃO, fornecidos pelo programa STEP.COMP.

Residuo	r_j	$F(r_j)$	$S(r_j)$	$ F(r_j) - S(r_j) $	
-	2.37541	-1.710	0.04355	0.02777	0.04355
-	2.33471	-1.681	0.04632	0.05555	0.01855
-	2.09414	-1.508	0.06574	0.08333	0.01759
-	1.66126	-1.196	0.11575	0.11111	0.03241
-	1.63531	-1.177	0.11943	0.13888	0.01945
-	1.13136	-.814	0.20758	0.16666	0.06869
-	1.11891	-.805	0.21017	0.19444	0.04351
-	1.08432	-.780	0.21741	0.22222	0.02296
-	1.02937	-.741	0.22922	0.25000	0.02077
-	.95668	-.689	0.24540	0.27777	0.03237
-	.90760	-.653	0.25665	0.30555	0.04889
-	.88271	-.635	0.26246	0.33333	0.07086
-	.76781	-.553	0.29012	0.36111	0.07098 ←
-	.62914	-.453	0.32522	0.38888	0.06366
-	.31542	-.227	0.41014	0.41666	0.02125
-	.20061	-.144	0.44255	0.44444	0.02589
-	.14097	-.101	0.45956	0.47222	0.01511
-	.07284	-.052	0.47907	0.50000	0.02092
+	.13249	+.095	0.53801	0.52777	0.03801
+	.14921	+.107	0.54279	0.55555	0.01501
+	.15138	+.109	0.54341	0.58333	0.03992
+	.36861	+.265	0.60468	0.61111	0.02134
+	.53105	+.382	0.64894	0.63888	0.03783
+	.57685	+.415	0.66109	0.66666	0.02221
+	.58253	+.419	0.66259	0.69444	0.03184
+	.62113	+.447	0.67269	0.72222	0.04952
+	.87014	+.626	0.73457	0.75000	0.01542
+	.89851	+.647	0.74122	0.77777	0.03655
+	.96994	+.698	0.75759	0.80555	0.04796
+	1.19220	+.858	0.80473	0.83333	0.02859
+	1.45432	+1.047	0.85255	0.86111	0.01921
+	1.57398	+1.133	0.87152	0.88888	0.01736
+	1.81053	+1.303	0.90388	0.91666	0.01499
+	1.92861	+1.389	0.91759	0.94444	0.02685
+	2.68841	+1.936	0.97358	0.97222	0.02913
+	2.83859	+2.044	0.97954	0.99999	0.02045

D = 0,0709

3. Calcula os autovalores $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k$ e seus respectivos autovetores.
4. Obtém e arquiva a matriz ortogonal de escores a ser utilizada na etapa seguinte.

Os componentes principais obtidos nesta primeira etapa são apresentados na Tabela 19, e os autovalores, autovetores e as matrizes ortogonais de escores nas Tabelas 20, 21 e 22.

A segunda etapa consiste em utilizar as matrizes de escores como novos conjuntos de variáveis independentes, agora denominadas COMP1, COMP2, ..., COMP8, processando, através do programa STEP.COMP novas análises de regressões.

São determinadas, nesta etapa, as porcentagens de explicabilidade de cada componente, apresentadas na Tabela 23.

Nos três cultivares, o primeiro componente foi suficiente para explicar a maior parte das variações na PPES (acima de 89% para C-111 e FLINT, e acima de 84% para o PIRANÃO). No entanto, os modelos foram determinados pelo programa STEPWISE, que incluiu todos os componentes cujo valor de t para os respectivos parâmetros fossem significativos a 5%. Os resultados são apresentados nas Tabelas 24, 25 e 26, as quais também apresentam os resíduos estimados e ordenados em relação a \bar{Y}_j .

Estes resíduos, quando estudados graficamente, apresentaram as mesmas tendências daqueles obtidos pelas análises de regressão usuais, fazendo-nos crer não ser a multicolinearidade a fon

te dos distúrbios observados, sendo também inadequados os ajustamentos pelos componentes principais (Figura 5).

Os resultados dos testes dos chorrilhos para os resíduos provenientes dos modelos de componentes principais também foram equivalentes aos anteriormente apresentados, indicando autocorrelação para os cultivares C-111 e PIRANÃO com as probabilidades $P(r \leq 12) < 0,015$ e $P(r \leq 13) < 0,031$, também inferiores ao nível $\alpha = 0,1$ estabelecido. Já para o cultivar FLINT, o teste não indicou autocorrelação, pois $P(r \leq 15) \leq 0,13$, certamente pelo maior número de parâmetros significativos presentes no modelo.

As distribuições nulas para estes testes são apresentadas na Tabela 27.

Os resultados obtidos para o teste de Lilliefors, e apresentados nas Tabelas 28, 29 e 30 para a normalidade dos resíduos, indicaram valores $D < 0,1477$, não havendo, portanto, motivos para não considerarmos os resíduos normalmente distribuídos.

Tabela 19 - Componentes principais para os três cultivares de milho.

C-111	Z ₁	=	0,42286	NINS	+	0,39646	TARM	-	0,44473	CARB	+	0,37926	UMID	+	0,28232	MASE	+	0,34771	PROT	+	0,05914	OLEO	+	0,34376	IACD
	Z ₂	=	0,07591	NINS	-	0,13641	TARM	-	0,04616	CARB	+	0,00324	UMID	-	0,59365	MASE	+	0,15118	PROT	+	0,74521	OLEO	+	0,20703	IACD
	Z ₃	=	-0,12991	NINS	-	0,13641	TARM	-	0,05251	CARB	-	0,15543	UMID	-	0,23064	MASE	+	0,51341	PROT	-	0,48691	OLEO	+	0,48657	IACD
	Z ₄	=	0,02292	NINS	-	0,12566	TARM	+	0,18826	CARB	+	0,68956	UMID	-	0,24724	MASE	-	0,50002	PROT	-	0,20779	OLEO	+	0,34410	IACD
	Z ₅	=	0,42672	NINS	-	0,11381	TARM	+	0,24487	CARB	-	0,47676	UMID	+	0,30744	MASE	-	0,35302	PROT	+	0,12149	OLEO	+	0,53269	IACD
	Z ₆	=	0,46724	NINS	+	0,49609	TARM	+	0,28101	CARB	-	0,18461	UMID	-	0,53484	MASE	+	0,05647	PROT	-	0,33729	OLEO	-	0,13966	IACD
	Z ₇	=	-0,63044	NINS	+	0,61949	TARM	+	0,09768	CARB	-	0,14412	UMID	-	0,02841	MASE	-	0,06377	PROT	+	0,05661	OLEO	+	0,42470	IACD
	Z ₈	=	0,00799	NINS	+	0,00032	TARM	+	0,78320	CARB	+	0,27265	UMID	+	0,26860	MASE	+	0,45905	PROT	+	0,17078	OLEO	-	0,01209	IACD
FLINT	Z ₁	=	-0,40534	NINS	-	0,39847	TARM	+	0,43887	CARB	-	0,42798	UMID	-	0,14938	MASE	-	0,40010	PROT	+	0,24292	OLEO	-	0,24437	IACD
	Z ₂	=	0,07158	NINS	+	0,26284	TARM	-	0,16504	CARB	+	0,10025	UMID	+	0,60514	MASE	-	0,28185	PROT	+	0,53457	OLEO	-	0,39634	IACD
	Z ₃	=	0,43009	NINS	+	0,30637	TARM	+	0,11073	CARB	-	0,24266	UMID	-	0,40787	MASE	-	0,09939	PROT	+	0,56848	OLEO	+	0,38806	IACD
	Z ₄	=	0,04247	NINS	-	0,15124	TARM	+	0,17628	CARB	-	0,14195	UMID	+	0,58636	MASE	-	0,27891	PROT	-	0,14698	OLEO	+	0,69344	IACD
	Z ₅	=	-0,26046	NINS	-	0,28852	TARM	-	0,37069	CARB	-	0,36794	UMID	+	0,16583	MASE	+	0,61242	PROT	+	0,38481	OLEO	+	0,15960	IACD
	Z ₆	=	-0,32024	NINS	+	0,72476	TARM	+	0,17182	CARB	-	0,49124	UMID	+	0,11079	MASE	+	0,15455	PROT	-	0,25009	OLEO	-	0,05107	IACD
	Z ₇	=	0,68786	NINS	-	0,21584	TARM	+	0,15179	CARB	-	0,46189	UMID	+	0,18584	MASE	+	0,18768	PROT	-	0,22462	OLEO	-	0,35162	IACD
	Z ₈	=	-0,01533	NINS	+	0,01659	TARM	+	0,73947	CARB	+	0,37107	UMID	+	0,16424	MASE	+	0,48848	PROT	+	0,22217	OLEO	-	0,00278	IACD
PIRANÃO	Z ₁	=	-0,35269	NINS	-	0,34420	TARM	+	0,37734	CARB	-	0,40608	UMID	-	0,33925	MASE	-	0,31174	PROT	+	0,29354	OLEO	-	0,38907	IACD
	Z ₂	=	0,18426	NINS	+	0,27337	TARM	-	0,31442	CARB	-	0,03288	UMID	-	0,39144	MASE	+	0,43810	PROT	+	0,63375	OLEO	-	0,21107	IACD
	Z ₃	=	0,53442	NINS	+	0,52172	TARM	+	0,34943	CARB	+	0,05743	UMID	-	0,12897	MASE	-	0,53899	PROT	+	0,06406	OLEO	-	0,07740	IACD
	Z ₄	=	-0,22450	NINS	+	0,22361	TARM	+	0,20666	CARB	+	0,02772	UMID	+	0,71799	MASE	-	0,21043	PROT	+	0,30693	OLEO	-	0,44956	IACD
	Z ₅	=	0,42982	NINS	-	0,24437	TARM	+	0,00797	CARB	-	0,63391	UMID	+	0,35952	MASE	+	0,02271	PROT	+	0,28392	OLEO	+	0,37845	IACD
	Z ₆	=	0,06034	NINS	-	0,44444	TARM	-	0,10064	CARB	+	0,58290	UMID	-	0,01874	MASE	-	0,35239	PROT	+	0,48562	OLEO	+	0,29758	IACD
	Z ₇	=	-0,56263	NINS	+	0,47601	TARM	+	0,03435	CARB	-	0,14659	UMID	-	0,06391	MASE	-	0,13221	PROT	+	0,22054	OLEO	+	0,60329	IACD
	Z ₈	=	-0,02638	NINS	+	0,02700	TARM	+	0,76325	CARB	+	0,25851	UMID	+	0,25657	MASE	+	0,48363	PROT	+	0,22087	OLEO	+	0,02586	IACD

Tabela 20 - Autovalores, autovetores e matriz ortogonal de escores do cultivar C-111, fornecidos pelo programa AUTO-JACOBI.

AUTO VALORES

$\lambda 1 = + 4.76498$
 $\lambda 2 = + 1.29758$
 $\lambda 3 = + .91669$
 $\lambda 4 = + .42791$
 $\lambda 5 = + .35884$
 $\lambda 6 = + .18870$
 $\lambda 7 = + .04486$
 $\lambda 8 = + .00040$

AUTO VETORES NAS LINHAS

+ .42296 + .39646 - .44473 + .37926 + .28232 + .34771 + .05914 + .34376
+ .07591 - .13641 - .04616 + .00324 - .59365 + .15118 + .74521 + .20703
- .12991 - .40690 - .05251 - .15543 - .23064 + .51341 - .48691 + .48657
+ .02292 - .12566 + .18826 + .68956 - .24724 - .50002 - .20779 + .34410
+ .42672 - .11381 + .24487 - .47676 + .30744 - .35302 + .12149 + .53269
+ .46724 + .49609 + .28101 - .18461 - .53484 + .05647 - .33729 - .13966
- .63044 + .61949 + .09768 - .14412 - .02841 - .06377 + .05661 + .42470
+ .00799 + .00032 + .78320 + .27265 + .26860 + .45905 + .17078 - .01209

NUMERO DE ROTACOES= 98

MATRIZ ORTOGONAL

- 2.48903 + .28849 + 1.19849 - 1.17188 - .19328 - .15475 + .03577 - .00881
- 3.06301 + .39688 + .17845 + .00118 + .28455 - .25683 + .16431 - .00613
- 2.80018 - .32619 + 1.21845 - .45039 - .21755 + .16675 - .00291 - .00086
- .55356 + .39189 + .68946 - .92044 - 1.02136 + .21965 - .19460 + .01405
- 1.64773 + .01404 + .56831 - .72570 - .01203 + .71725 + .17333 + .00164
- 1.78528 + .09335 + .11949 - .62933 + .09300 + .30112 + .35226 - .00413
- .05638 + .89614 + .13335 - .08031 - .16598 + .20524 - .42787 - .04756
- 1.30612 + .73157 - .56022 + .33696 + .36459 + .31221 - .02343 + .00752
- 1.32239 + .21174 - .77745 + .90090 + .90165 + .59136 - .28065 + .01158
- .36832 + 1.24805 - .30027 - .22999 + .23636 + .16685 - .36129 + .00877
- .71500 + 1.00082 - .81607 + .73926 + .15301 + .02165 - .23507 + .01503
- .77780 + .42780 - .11189 + .34874 + .30517 + .57436 - .35242 + .01490
- 2.61337 - .45825 + .31107 + .75191 + .04844 - .46365 + .05005 - .00077
- 2.82089 - .24692 - .07851 + .66130 + .49536 - .64421 + .10183 - .00950
- 2.86227 - .48082 + .25477 + .67432 - .11140 - .66735 - .03457 + .00089
+ .45798 + .43424 - .34869 - .02111 - .72504 - .20263 + .35113 + .01095
+ .22303 - .12619 - .56570 + .83328 - .35365 + .24850 + .11581 + .01897
- .28943 - .00424 - .21836 + .05695 - .62078 + .17461 + .55141 + .01002
+ 1.89508 + 1.15232 - .27929 - .02938 - .32049 - .45677 - .08870 + .01072
+ 2.89531 + 1.39706 + .64522 + .04856 - .26148 - .37960 - .10285 + .00967
+ 2.34009 + 1.96811 - .39661 - .09376 + .52053 - .65423 + .02865 - .00287
+ 1.33242 + 1.77388 - .58058 - 1.02152 + .05584 - .36848 + .00754 - .00018
+ 3.03731 + 1.03302 + 1.38804 + .65925 + 1.43156 + .13947 + .23326 - .00991
+ 4.86381 + .24953 + 3.39928 + .92181 - .03735 + .01536 + .09560 + .00583
- 3.13794 - 1.04886 + .79484 - .31351 + .19360 + .05339 - .03004 - .00655
- 2.46391 - .45573 + .68693 - .20326 - .43070 - .33012 - .18688 + .00237
- 2.96615 - .71574 + .05425 + .50629 + .04759 - .35038 - .09025 + .00020
+ 1.54675 + .11185 - .90993 + .90271 - 1.24863 + .46651 + .06073 - .08885
+ 2.03994 + .62934 - .39903 - .87288 - 1.06367 + .37446 - .05791 + .02254
+ .61377 + .35384 - 1.50627 + .61468 - .28572 + .69000 + .25869 + .02406
+ 2.40555 - 3.70895 + .12844 + .84421 - .95389 - .04735 - .19448 + .02330
+ 2.33439 - 1.14192 - 1.87021 - .49188 - .01216 - 1.25965 - .02361 - .00492
+ 1.76149 - .69255 - .98801 - .61696 + .30037 + .14507 - .07144 + .00352
+ 2.69660 - 2.05909 + .17869 - .32633 + .55570 + .11247 - .07252 - .00279
+ .66101 - .98133 - 1.57736 - .35199 + 1.05478 + .33866 + .21948 - .00603
+ 2.43423 - 2.35558 + .33693 - 1.25167 + .99228 + .20599 - .04834 - .01669

Tabela 21 - Autovalores, autovetores e matriz ortogonal de escores do cultivar FLINT, fornecidos pelo programa AUTO-JACOBI.

AUTO VALORES

λ 1= + 4.65292
 λ 2= + 1.28269
 λ 3= + .93397
 λ 4= + .76141
 λ 5= + .21394
 λ 6= + .08873
 λ 7= + .06595
 λ 8= + .00036

AUTO VETORES NAS LINHAS

- .40534 - .39847 + .43887 .42798 .14938 - .40010 + .24292 - .24437
 + .07158 + .26284 .16504 + .10025 + .60514 - .28185 + .53457 - .39634
 + .43009 + .30637 + .11073 - .24266 .40787 - .09939 + .56848 + .38806
 + .04247 - .15124 + .17628 .14195 + .58636 - .27891 - .14698 + .69344
 - .26046 - .28852 - .37069 .36794 + .16583 + .61242 + .38481 + .15960
 - .32024 + .72476 + .17182 .49124 + .11079 + .15456 - .25009 - .05107
 + .68786 - .21584 + .15179 .46189 + .18584 + .18768 - .22462 - .35162
 - .01533 + .01659 + .73947 + .37107 + .16424 + .48848 + .22217 - .00278
 NUMERO DE ROTACOES= 106

MATRIZ ORTOGONAL

+ 2.92867 - 1.54539 + .31246 - .11839 + .46002 + .23853 - .09356 + .00036
 + 2.83369 1.35970 - .03972 - .24014 + .19397 + .13797 - .08866 + .00135
 + 2.97375 1.39817 + .13669 - .54222 + .24692 + .11340 - .08513 + .00298
 + 1.61542 .64352 + .18195 - .01492 + .19128 + .69025 + .10894 + .00814
 + 1.04111 .32006 + .16444 - .22828 + .10040 + .24446 + .31182 + .00184
 + .86795 .41277 + .18121 - .20265 + .08550 + .11454 + .10813 + .00197
 + .01438 .32047 + 1.17483 + .51525 + .67667 - .46175 - .02273 - .01205
 + .92983 .41823 + 1.01729 + .58528 + .45139 + .08068 + .01859 - .00509
 + .59603 1.15896 + .63131 + .77334 - .59635 - .07643 - .07407 - .00738
 + .36967 - .57057 + 1.95616 + .23635 - .17566 - 1.02294 - .42394 - .01489
 + .87006 .37911 + 1.17305 - .46109 + .06934 - .28554 + .41157 - .00834
 + .96914 .20547 + .91579 + .40659 - .13578 - .13184 + .25720 - .00896
 + 1.84194 + .55087 - 1.77375 + .43925 + .11482 - .24121 - .03618 - .00060
 + 2.20684 + .07454 - 1.37451 + .17057 + .06764 - .08587 + .04985 + .00069
 + 2.19870 + .27299 - 1.22992 + .03036 + .39289 - .12370 + .04819 + .00042
 .12672 + 1.13229 + .33446 + .51010 - .25022 + .13546 + .35849 - .00192
 .51270 + 1.00989 - .20339 - .47448 + .22455 + .04477 - .08961 + .01205
 + .40110 + .18877 + .18359 - .49758 - .05720 + .39307 - .66407 + .02260
 .41758 + .18097 + .95673 + 1.45946 - .36100 + .13693 + .10761 - .00781
 .72621 + .44423 + .26672 + 1.65740 + .71200 + .47981 - .40221 + .00442
 1.03507 + 1.19236 + .01813 + 2.66999 - .16161 + .01443 - .37758 + .01297
 1.43326 + 1.31533 - .23480 + .84646 + .40016 - .06910 + .39310 - .00518
 + .05833 + .05555 + .76171 + .02302 - .68115 + .13167 + .21696 + .00140
 .80017 + .15574 + .27897 + .21531 - .16409 + .17189 + .54742 - .00333
 + 2.36650 .22708 - 1.71192 + .03257 - .75295 - .09224 + .08116 + .00216
 + 1.65355 + .15698 - 1.48525 - .86900 + .65669 - .39156 - .07049 + .00554
 + 2.35840 + .00239 - 1.73814 - .27946 - .86793 - .32256 - .06256 + .00331
 3.03916 + 2.69516 - 1.05545 + .27944 + .42696 - .21734 + .03374 + .00844
 1.45563 + 1.12359 - .54856 - 1.16414 - .10705 + .39021 - .32666 - .09824
 1.93645 + 1.62029 - .09539 - .91191 + .09954 + .00423 + .06885 + .01536
 1.34303 + .71371 + .39816 - .57869 - .91877 + .00207 - .30954 + .01580
 .65464 + .41985 + .82455 - 1.13394 - .96846 + .21784 - .15280 + .01831
 2.23115 + .46107 + .48362 - .86040 - .28117 - .09837 + .15090 + .00820
 3.74980 - 1.07063 + .21073 - 1.40826 + .61060 - .13042 - .21341 + .01427
 2.05830 + .31115 + .86321 - 1.65888 + .47764 + .02157 + .13644 + .01428
 7.57522 4.04762 - 1.93500 + .79370 - .17962 - .01294 + .08423 - .00313

Tabela 22 - Autovalores, autovetores e matriz ortogonal de escores do cultivar PIRANÃO, fornecidos pelo programa AUTO-JACOBI.

AUTO VALORES

$\lambda 1 = + 5.45582$
 $\lambda 2 = + 1.13609$
 $\lambda 3 = + .76444$
 $\lambda 4 = + .32163$
 $\lambda 5 = + .14597$
 $\lambda 6 = + .10576$
 $\lambda 7 = + .06584$
 $\lambda 8 = + .00442$

AUTO VETORES NAS LINHAS

- .35269 - .34420 + .37734 - .40608 - .33925 - .31174 + .29354 - .38907
+ .18426 + .27338 - .31442 - .03288 - .39144 + .43810 + .63375 - .21107
+ .53442 + .52172 + .34943 + .05743 - .12897 - .53899 + .06486 - .07440
- .22450 + .22361 - .20666 + .02772 + .71799 - .21043 + .30693 - .44956
+ .42982 - .24437 + .00797 - .63391 + .35952 + .02271 + .28392 + .37845
+ .06034 - .44444 - .10064 + .58290 - .01874 - .35239 + .48562 + .29758
- .56263 + .47601 + .03435 - .14659 - .06391 - .13221 + .22054 + .60329
- .02638 + .02700 + .76325 + .25851 + .25657 + .48363 + .22087 + .02586
NUMERO DE ROTACOES= 98

MATRIZ ORTOGONAL

+ 2.76713 - .84817 - .49823 - .53073 + .26830 - .01966 + .46328 + .00236
+ 2.80195 - .29474 - .64577 - .35586 + .34176 + .19941 + .49225 + .00567
+ 2.77085 + .16964 - .91875 - .48382 + .21524 + .03184 + .34098 + .00519
+ 1.35087 + .83932 - .26921 - .21995 + .11935 - .18337 + .08626 - .00141
+ .71629 + 1.41054 - .01641 - .05603 + .15754 + .71639 - .19599 - .03379
+ 1.04719 + .46734 - .10091 - .22213 + .09607 - .14605 - .08239 - .01281
+ .75303 + .85144 - .01615 + .03660 + .31241 + .31176 .19109 - .02177
+ .98683 + .47647 + .20263 - .08499 + .11670 + .24838 - .14864 - .02842
+ 1.39895 + .52057 - .32513 - .20224 + .52823 - .66535 + .11289 + .02090
+ 1.41947 - .42220 + .91348 - .26391 - .23206 + .22105 - .24165 + .36543
+ .86620 + .42585 - .09071 - .26079 + .37639 - .41837 - .29294 - .01017
+ .39024 + .66066 - .30642 - .11186 + .09335 + .18777 - .18429 - .02016
+ 1.98885 - .91905 - .85526 - .04731 - .39947 + .12895 - .09672 - .03108
+ 2.02231 - 1.10356 - .75970 + .10573 - .38249 + .12826 - .13736 + .00836
+ 2.07016 - 1.37048 - .69101 - .07440 - .45784 - .10992 - .19871 - .03566
+ .85547 - .25596 + 1.20342 + .14592 - .29085 + .17553 + .37365 - .02168
+ .62597 - .33284 + 1.59174 + .05181 + .01344 + .15665 + .01019 - .04156
+ .98123 - .31649 + 1.40835 + .05134 + .11678 - .09461 + .27607 - .01915
- .24597 + .64652 + 1.00300 - .24148 + .44779 + .29506 - .04840 - .02570
+ .14386 + .43530 + .94290 - .09796 + .26539 - .15917 - .16464 - .02346
+ .40601 - .23974 + .86294 - .02318 - .30870 - .42291 - .07566 - .02690
- .88004 + .95358 + .26242 - .32990 + .43246 - .06789 - .17847 + .00432
- 1.23801 - .53282 + .52551 - .77319 - .61853 + .37279 + .21062 - .03990
- .34378 + .62489 + .86634 - .41330 + .33372 - .39077 - .57463 - .03434
+ 1.90725 - 1.59411 - .61988 + .20485 - .41847 + .12406 - .15838 - .03833
+ .90797 - .49602 - 1.94741 + .77411 + .03418 - .14530 - .34549 + .00337
+ 1.97653 - 1.05494 - .89627 + .78310 + .20896 + .15514 - .03727 - .00741
- .19757 + .32290 + 1.17358 + .39142 - .50410 - .19981 + .41614 - .00935
- 3.60531 + 1.71807 - 1.13862 - .21648 - .61510 - .10849 - .00234 + .01194
- 2.03838 + 1.98371 - .93796 + .18036 - .21345 - .62497 + .29388 + .04497
- 1.53966 + .41755 + .46480 - .23276 - .64430 - .40828 + .19467 - .00789
- 1.44571 + .55997 + 1.13829 + 1.27608 - .14410 + .52472 - .13685 - .03183
- 2.00295 - .39544 + .05342 + 2.37566 + .37569 - .25009 + .22077 + .03957
- 7.32186 - 2.06956 - .65532 - .49806 + .94996 + .33215 + .30186 + .03110
- 4.84078 + 1.72028 - 1.38070 - .14717 - .40453 + .50534 + .01720 + .00867
- 5.45463 - 2.95848 + .45705 - .48941 - .16975 - .40323 - .31873 - .02901

Tabela 23 - Resultados das análises de componentes principais nos três cultivares de milho.

COMPONENTES PRINCIPAIS	Autovalores			% de variância explicada			% de variância acumulada			% de explicabilidade		
	C-111	FLINT	PIRANÃO	C-111	FLINT	PIRANÃO	C-111	FLINT	PIRANÃO	C-111	FLINT	PIRANÃO
COMP1	4,76498	4,65292	5,45582	59,56	58,17	68,20	59,56	58,17	68,20	89,45*	89,17*	84,15*
COMP2	1,29758	1,28269	1,13609	16,22	16,03	14,20	75,78	74,20	82,40	0,10	1,14*	2,10*
COMP3	0,91669	0,93397	0,76444	11,46	11,69	9,56	87,24	85,89	91,95	2,01*	1,26*	9,40*
COMP4	0,42791	0,76141	0,32163	5,35	9,52	4,02	92,59	95,41	95,98	0,00	1,15*	0,20
COMP5	0,35884	0,21394	0,14597	4,49	2,67	1,83	97,08	98,08	97,80	4,45*	2,64*	0,00
COMP6	0,18870	0,08873	0,10576	2,36	1,12	1,32	99,44	99,20	99,11	0,62*	0,72*	0,20
COMP7	0,04486	0,06595	0,06584	0,56	0,08	0,08	100,00	100,00	100,00	0,38	0,52*	0,08
COMP8	0,00040	0,00036	0,00442	0,00	0,00	0,00	100,00	100,00	100,00	0,11	0,55*	0,14
R ²										97,12	97,15	96,27

* Componentes selecionados pela regressão STEPWISE.

Tabela 24 - Resultados da análise de regressão dos componentes principais selecionados para o cultivar C-111.

Comp. pr.	Média	Desvio Padrão	Autovvalor
COMP1	.00000	2.18288	4.76
COMP2	.00000	1.13911	1.29
COMP3	.00000	.95744	.91
COMP4	.00000	.65415	.42
COMP5	.00000	.59303	.35
COMP6	.00000	.43440	.18
COMP7	.00000	.21181	.04
COMP8	.00000	.02318	.00
PPES	16.66330	9.13549	83.45

Matriz de Correlações

+1.0000								
+0.0000	+1.0000							
-0.0000	-0.0000	+1.0000						
-0.0000	+0.0000	-0.0000	+1.0000					
-0.0000	+0.0000	-0.0000	+0.0000	+1.0000				
+0.0000	-0.0000	+0.0000	+0.0000	+0.0000	+1.0000			
+0.0000	+0.0000	+0.0000	+0.0000	+0.0000	+0.0000	+1.0000		
+0.0000	+0.0000	+0.0000	+0.0000	+0.0000	+0.0000	+0.0000	+1.0000	
+0.0000	+0.0000	+0.0000	+0.0000	+0.0000	+0.0000	+0.0000	+0.0000	+1.0000

Regressão Passo a Passo

Resíduo Padrão Residual= 1.8044796
 Coeficiente de Correlação= .988569689
 A. de Variacao Explicada= 96.5443195
 erro Constante= 16.6633073

Análise da Variância

C. Variacao	G.L.	S.O.	G.M.	F.
Regressão	4	2620.0627	705.01568	216.51
Resíduo	31	100.2405	3.23514	
Total	35	2921.0032		

Comp. pr.	Coeficiente B	Erro Padrão	Intervalo de Confiança
COMP1	3.955472	*1 .13972	3.87656 C=B 1C= 4.24838
COMP3	1.850802	*1 .31857	2.00036 C=B 3C= 4.70123-
COMP5	3.219465	*1 .50317	2.18126 C=B 5C= 4.25766
COMP6	1.854751	*1 .70210	.22310 C=B 6C= 3.08642

VALORES ORDENADOS EM RELAÇÃO A Y. est.

Obs	Y. obs.	Y. est.	Resíduo/s	Cls	Resíduo
3	5.4255	3.5144	+ 1.05904	3	1.91132
25	4.0398	3.8898	+ 0.6874	25	1.1504
26	4.8878	4.0551	+ 0.46147	26	0.8321
1	4.3601	4.3192	+ 0.2257	1	0.491
27	5.2397	4.4280	+ 0.4934	27	0.8173
2	5.4576	4.7944	+ 0.5543	2	0.6547
13	6.8122	5.9928	+ 0.8193	13	1.5142
15	7.9370	5.5111	+ 1.36895	15	2.4187
14	6.8324	6.1377	- 0.2743	14	0.4517
6	7.5497	10.2411	- 1.7313	6	2.5217
5	8.1484	10.5189	- 1.81665	5	2.3025
4	9.4185	10.6219	- 0.6095	4	1.2034
8	12.4856	13.9462	- 0.8045	8	1.4364
18	9.8423	14.1089	- 2.36439	18	4.2651
11	14.5863	15.4697	- 0.8734	11	0.874
12	16.1189	15.6744	+ 0.4632	12	0.443
7	15.3738	16.0712	- 0.6974	7	1.0074
16	13.6448	16.2806	- 1.46039	16	2.6335
9	14.2405	16.3582	- 0.79035	9	1.4393
10	17.2765	16.6559	+ 0.4583	10	0.621
17	17.4895	17.8889	- 0.6793	17	1.009
28	23.0850	20.0733	+ 1.2311	28	2.9174
20	22.7471	21.3555	+ 1.1166	20	1.3953
22	22.6401	22.2579	+ 0.3821	22	0.4345
29	24.8367	22.4785	+ 1.29032	29	2.6219
19	22.9911	22.7635	+ 0.2276	19	0.2619
31	22.3173	23.0637	- 0.6554	31	0.9095
35	24.6461	23.3123	+ 0.6371	35	1.3321
20	25.6141	23.7867	+ 1.8146	20	3.8271
33	30.1479	26.1838	+ 2.19580	33	3.9333
32	24.0381	26.3125	- 1.93214	32	3.2333
21	25.7163	27.0615	- 0.7446	21	1.3410
34	31.0953	29.0715	+ 1.1166	34	2.0267
36	30.4327	29.3854	+ 0.6036	36	1.0473
24	30.3612	31.2359	- 0.4945	24	0.8133
23	31.5022	31.5571	+ 0.9552	23	1.1513

Tabela 25 - Resultados da análise de regressão dos componentes selecionados para o cultivar FLINT.

Comp. nr.	Medida	Desvio. Padrão	Autocorrel.
COMP1	.00000	2.15706	4.55
COMP2	.00000	1.13255	1.28
COMP3	.00000	.86642	.93
COMP4	.00000	.87359	.76
COMP5	.00000	.46253	.21
COMP6	.00000	.29788	.08
COMP7	.00000	.25681	.06
COMP8	.00000	.01905	.00
FPES	14.05412	6.66125	47.35

Matriz de Correlações

+1.0000							
+0.0000	+1.0000						
+0.0000	-0.0000	+1.0000					
-0.0000	-0.0000	-0.0000	+1.0000				
-0.0000	-0.0000	-0.0000	-0.0000	+1.0000			
+0.0000	-0.0000	+0.0000	-0.0000	-0.0000	+1.0000		
+0.0000	+0.0000	+0.0000	-0.0000	-0.0000	+0.0000	+1.0000	
-0.9443	+0.1067	-0.1071	-0.1623	-0.0948	+0.0716	+0.0742	+1.0000

Regressão Passo a Passo

Desvio Padrão Residual= 1.32265004
 Coeficiente de Correlação= .98562405
 % de Variacao Explicada= 97.1454769
 Termo Constante= 14.05412

Análise da Variância

C. Variacao	G.L.	S.Q.	G.M.	F
Regressão	8	1609.9281	201.24376	114.85
Resíduo	27	47.3881	1.75511	
Total	35	1657.3162		

Comp. nr.	Coefficiente B	Erro Padrão	Intervalo de Confiança
COMP1	3.012603- *	.10372	3.22534-C=B 1C= 2.79936-
COMP2	.548408 *	.19755	1.24322-C=B 2C= 1.05358
COMP3	.726757 *	.23151	1.32391-C=B 3C= 1.27360
COMP4	.844941- *	.25641	1.37084-C=B 4C= .31903-
COMP5	2.415116- *	.48373	3.40724-C=B 5C= 1.42238-
COMP6	1.250037- *	.75110	3.50054-C=B 6C= .41552-
COMP7	1.919962 *	.87122	1.13308-C=B 7C= 3.78634
COMP8	26.836618 *	11.74226	2.75023-C=B 8C= 50.91698

VALORES ORDEMADOS EM RELACAO A Y.est.

Obs	Y. obs.	Y. est.	Resíduo/s
1	4.3146	2.8392	1.4754
2	4.5921	3.5540	1.0381
3	5.3673	3.9541	1.4132
15	7.6366	5.9956	1.6410
14	7.3214	6.3313	1.0001
13	8.1770	7.1340	1.0430
4	6.3586	7.5410	-1.1824
25	6.4475	7.5958	-1.1483
26	7.1564	7.9173	-0.7609
27	7.5261	8.4358	-0.9097
8	10.2165	9.2503	0.9662
5	10.8518	10.9408	0.9110
6	9.3931	11.3173	-1.9242
20	11.9235	12.8289	-0.9054
12	13.1553	12.3258	0.8295
18	10.4880	12.2343	-1.7463
10	10.5048	12.6974	-2.1926
7	11.8207	13.2826	-1.4619
11	12.3336	13.4723	-1.1387
10	14.5900	15.1483	-0.5583
19	16.8327	15.5959	1.2368
21	16.3420	15.6892	0.6528
29	16.0414	15.9437	0.0977
16	16.6743	15.3817	1.2926
17	15.1508	16.0132	-0.8624
23	15.8063	16.3446	-0.5383
24	16.6472	17.5274	-0.8802
22	17.5325	18.1058	-0.5733
32	19.3825	20.0253	-0.6428
31	20.7729	21.4146	-0.6417
30	21.3326	21.5288	-0.1962
28	22.8924	21.5373	1.3551
28	22.5041	23.1565	-0.6524
33	26.1154	23.5697	2.5457
34	24.6756	24.7590	0.9166
36	32.5518	32.5712	0.9806

Tabela 26 - Resultados da análise de regressão dos componentes selecionados para o cultivar PIRANAO.

Comp. nr.	Media	Desvio Padrao	Autovvalor
COMP1	.00000	2.33577	5.45
COMP2	.00000	1.06587	1.13
COMP3	.00000	.87432	.76
COMP4	.00000	.56712	.32
COMP5	.00000	.38386	.14
COMP6	.00000	.32521	.10
COMP7	.00000	.26559	.06
COMP8	.00000	.06556	.00
PPES	14.34203	7.31439	53.50

Matriz de Correlacoes

+1.0000					
+0.0000	+1.0000				
+0.0000	-0.0000	+1.0000			
-0.0000	+0.0000	-0.0000	+1.0000		
+0.0000	+0.0000	+0.0000	-0.0000	+1.0000	
+0.0000	-0.0000	-0.0000	+0.0000	-0.0000	+1.0000
+0.0002	-0.0000	+0.0000	-0.0000	+0.0000	-0.0000
-0.9173	+0.1449	+0.3065	-0.0438	-0.0072	-0.0273

Regressao Passo a Passo

Desvio Padrao Residual= 1.55519909
 Coeficiente de Correlacao= .978013056
 % de Variacao Explicada= 95.6513451
 Termo Constante= 14.3420384

Analise da Variacao

C. Variacao	G.L.	S.Q.	Q.M.	F
Regressao	3	1731.0937	597.02791*	234.61
Residuo	32	81.4291	2.54466	
Total	35	1872.5128		

Comp. nr.	Coefficiente B	Erro Padrao	Intervalo de Confianca
COMP1	2.872712-*	.11543	3.10774- C= B 1C= 2.63767-
COMP2	.954585 *	.25297	.47953 C= B 2C= 1.50963
COMP3	2.564109 *	.30339	1.93621 C= B 3C= 3.19200

VALORES ORDENADOS EM RELACAO A Y. est.

Obs	Y. obs.	Y. est.	Residuo's
3	4.3681	4.1951	+ .1830
1	4.7613	4.2717	+ .4896
2	4.1312	4.3433	- .2121
15	7.1597	5.2601	+ 1.8996
27	5.7247	5.3165	+ .4082
14	6.8825	5.4865	+ 1.3960
13	6.9710	5.5215	+ 1.4495
25	5.8021	5.0715	+ .7306
26	6.0831	5.2459	+ .8372
9	8.7306	10.0073	- 1.2767
4	10.1672	10.6868	- .5196
6	10.1209	11.5338	- 1.4129
11	11.2634	12.0446	- .7812
10	11.1552	12.1866	- 1.0314
8	11.7623	12.5076	- .7453
7	11.3243	12.9343	- 1.6100
12	11.9532	13.0933	- 1.1401
5	12.7633	13.6451	- .8818
16	14.4857	14.7156	- 1.2299
18	17.2933	14.8156	+ 2.4777
21	12.6448	15.1459	- 2.5011
17	15.4337	16.2911	- .8574
20	17.2482	16.7754	+ .4728
24	21.5643	18.1753	+ 3.3890
28	18.0055	18.2633	- .2578
19	20.4181	18.2634	+ 2.1547
22	19.9659	18.4914	+ 1.4745
23	17.4275	18.7160	- 1.2885
30	17.1335	19.7656	- 2.6321
33	17.0224	19.8236	- 2.8012
31	21.9397	20.3721	+ 1.5676
32	23.0015	21.9708	+ 1.0307
35	23.4897	23.4832	+ .0065
34	23.2714	26.4163	- 3.1449
36	26.6601	26.2411	+ .4190
34	32.1639	31.6369	+ .5270

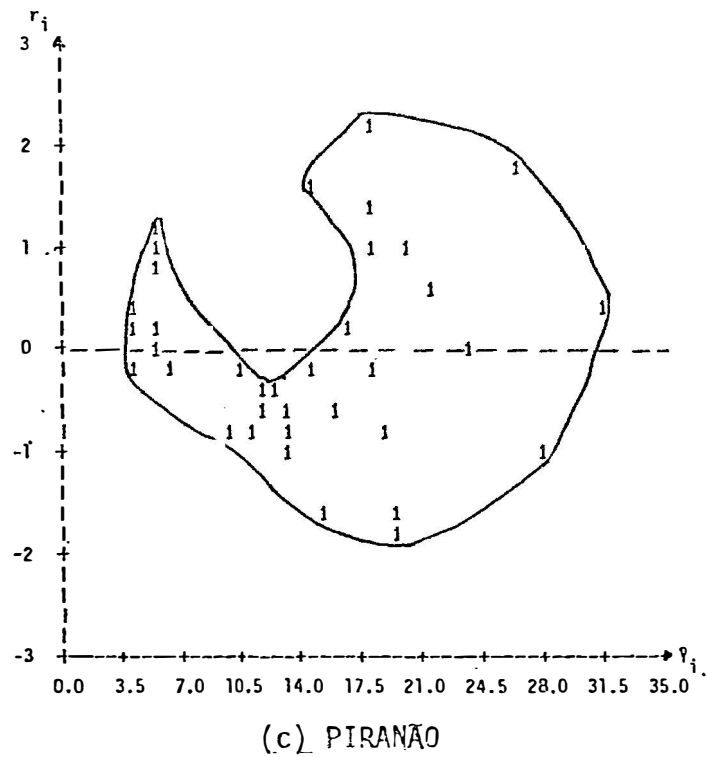
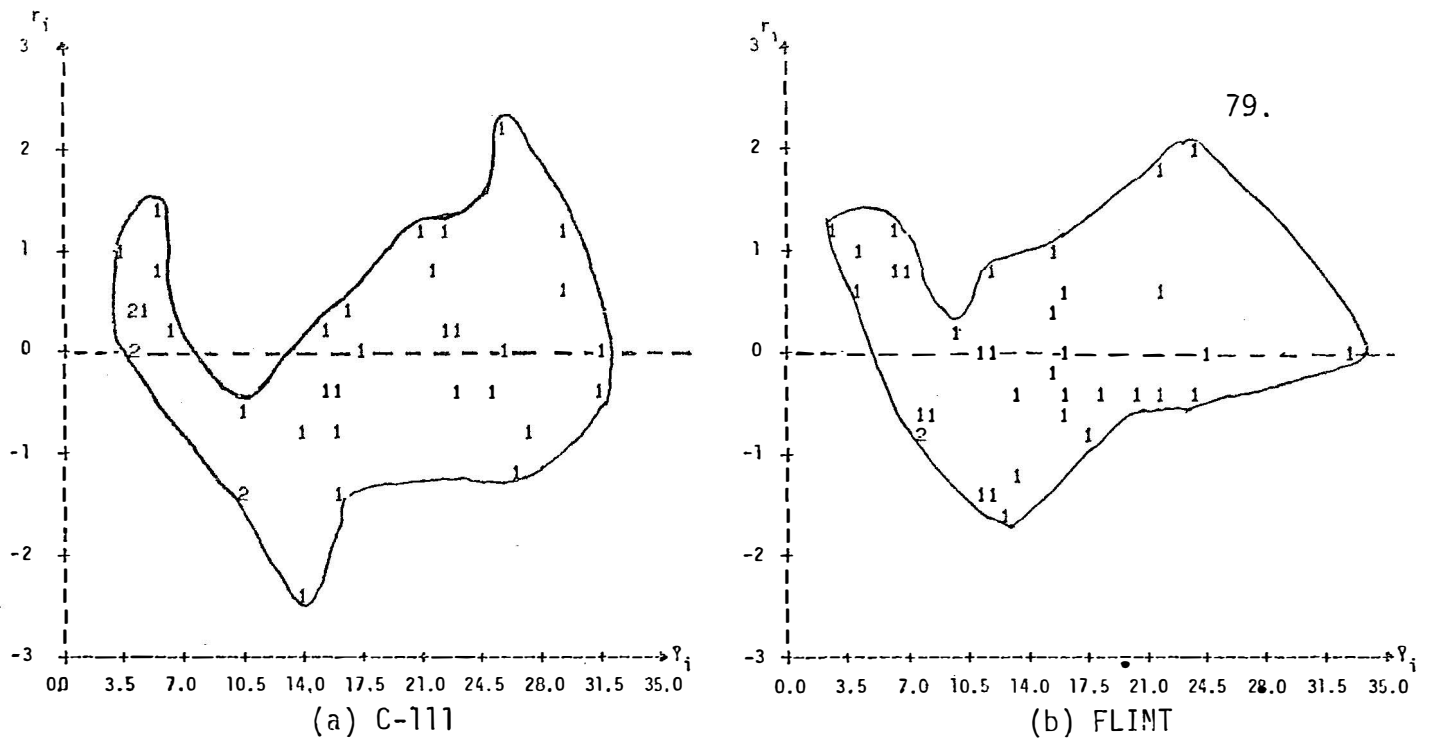


Figura 5 - Gráficos dos resíduos obtidos dos modelos de componentes principais.

Tabela 28 - Resultados do teste de Lilliefors para normalidade dos resíduos do cultivar C-111, provenientes do modelo de componentes principais.

	Residuo	r_i	$F(r_i)$	$S(r_i)$	$ F(r_i)-S(r_i) $
-	4.26651	-2.364	0.00902	0.02777	0.01874
-	2.63886	-1.462	0.07181	0.05555	0.04403
-	2.59147	-1.436	0.07548	0.08333	0.01992
-	2.37048	-1.313	0.09447	0.11111	0.01663
-	2.22446	-1.232	0.10883	0.13888	0.03005
-	1.46064	-.809	0.20912	0.16666	0.07023
-	1.42563	-.790	0.21474	0.19444	0.04800
-	1.34319	-.744	0.22832	0.22222	0.03388
-	1.20334	-.666	0.25242	0.25000	0.03020
-	.88374	-.489	0.31215	0.27777	0.06215
-	.87473	-.484	0.31392	0.30555	0.03614
-	.72669	-.402	0.34357	0.33333	0.03802
-	.70034	-.388	0.34896	0.36111	0.01563
-	.55135	-.305	0.37997	0.38888	0.01886
-	.15486	-.085	0.46580	0.41666	0.07691
-	.10436	-.057	0.47693	0.44444	0.06027
+	.02541	+.014	0.50561	0.47222	0.06117
+	.04091	+.022	0.50904	0.50000	0.03682
+	.12404	+.068	0.52740	0.52777	0.02740
+	.23063	+.127	0.55085	0.55555	0.02307
+	.34216	+.189	0.57519	0.58333	0.01964
+	.44448	+.246	0.59728	0.61111	0.01395
+	.49467	+.274	0.60801	0.63888	0.03087
+	.62260	+.345	0.63496	0.66666	0.03170
+	.64317	+.356	0.63924	0.69444	0.05520
+	.81173	+.449	0.67358	0.72222	0.04863
+	.83271	+.461	0.67777	0.75000	0.07222
+	1.04730	+.580	0.71917	0.77777	0.05860
+	1.39163	+.771	0.77970	0.80555	0.02584
+	1.51942	+.842	0.80011	0.83333	0.03321
+	1.91102	+1.059	0.85520	0.86111	0.02187
+	2.01878	+1.118	0.86837	0.88888	0.02050
+	2.31174	+1.281	0.89992	0.91666	0.01674
+	2.32819	+1.290	0.90151	0.94444	0.04292
+	2.41594	+1.338	0.90969	0.97222	0.06253
+	3.96408	+2.196	0.98598	0.99999	0.01401

D = 0,0769

Tabela 29 - Resultados do teste de Lilliefors para normalidade dos resíduos do cultivar FLINT, provenientes do modelo de componentes principais.

	Residuo	r_i	$F(r_i)$	$S(r_i)$	$ F(r_i)-S(r_i) $
-	2.10266	-1.588	0.05608	0.02777	0.05608
-	1.92717	-1.455	0.07270	0.05555	0.04493
-	1.74554	-1.318	0.09363	0.08333	0.03807
-	1.52592	-1.152	0.12449	0.11111	0.04116
-	1.14828	-.867	0.19283	0.13888	0.08172
-	1.14238	-.863	0.19405	0.16666	0.05517
-	.98024	-.740	0.22948	0.19444	0.06281
-	.88968	-.672	0.25075	0.22222	0.05630
-	.86232	-.651	0.25737	0.25000	0.03515
-	.76089	-.574	0.28270	0.27777	0.03270
-	.66036	-.498	0.30893	0.30555	0.33115
-	.64367	-.486	0.31338	0.33333	0.01994
-	.64287	-.485	0.31360	0.36111	0.04750
-	.57427	-.433	0.33220	0.38888	0.05668
-	.53660	-.405	0.34259	0.41666	0.07406
-	.53503	-.404	0.34303	0.44444	0.10141
-	.15081	-.113	0.45464	0.47222	0.01757
-	.10627	-.080	0.46800	0.50000	0.03199
-	.09905	-.074	0.47017	0.52777	0.05760
-	.09347	-.070	0.47185	0.55555	0.08370
+	.09776	+.073	0.52943	0.58333	0.05389
+	.12056	+.091	0.53628	0.61111	0.07482
+	.26623	+.201	0.57970	0.63888	0.05918
+	.65279	+.493	0.68905	0.66666	0.05016
+	.69263	+.523	0.69960	0.69444	0.03293
+	.73807	+.557	0.71143	0.72222	0.01699
+	.85808	+.648	0.74158	0.75000	0.01936
+	.92800	+.701	0.75837	0.77777	0.01940
+	.99007	+.747	0.77275	0.80555	0.03279
+	.99301	+.750	0.77342	0.83333	0.05990
+	1.32171	+.998	0.84098	0.86111	0.02012
+	1.43325	+1.082	0.86054	0.88888	0.02834
+	1.48435	+1.121	0.86893	0.91666	0.04773
+	1.64001	+1.238	0.89232	0.94444	0.05212
+	2.36521	+1.786	0.96301	0.97222	0.01857
+	2.54576	+1.923	0.97277	0.99999	0.02722

D = 0,1014

Tabela 30 - Resultados do teste de Lilliefors para normalidade dos resíduos do cultivar PIRANÃO, provenientes do modelo de componentes principais.

	Residuo	r_i	$F(r_i)$	$S(r_i)$	$ F(r_i)-S(r_i) $
-	2.81722	-1.766	0.03869	0.02777	0.03869
-	2.63201	-1.649	0.04947	0.05555	0.02169
-	2.50510	-1.570	0.05816	0.08333	0.02517
-	1.58988	-.996	0.15946	0.11111	0.07612
-	1.58091	-.991	0.16083	0.13888	0.04972
-	1.41883	-.889	0.18688	0.16666	0.04799
-	1.28848	-.807	0.20962	0.19444	0.04295
-	1.27669	-.800	0.21175	0.22222	0.01731
-	1.13711	-.712	0.23797	0.25000	0.01575
-	1.03138	-.646	0.25895	0.27777	0.01881
-	.88284	-.553	0.28998	0.30555	0.01557
-	.85838	-.538	0.29525	0.33333	0.03808
-	.77519	-.485	0.31350	0.36111	0.04761
-	.73769	-.462	0.32188	0.38888	0.06700
-	.43860	-.274	0.39167	0.41666	0.02499
-	.23330	-.146	0.44186	0.44444	0.02519
-	.22984	-.144	0.44271	0.47222	0.02950
-	.21259	-.133	0.44698	0.50000	0.05301
-	.16380	-.102	0.45910	0.52777	0.06867
+	.00146	+.000	0.50036	0.55555	0.05518
+	.11407	+.071	0.52850	0.58333	0.05482
+	.16508	+.103	0.54121	0.61111	0.06989
+	.40810	+.255	0.60095	0.63888	0.03792
+	.46887	+.293	0.61559	0.66666	0.05107
+	.48962	+.306	0.62055	0.69444	0.07388
+	.52697	+.330	0.62943	0.72222	0.09278
+	1.03071	+.646	0.74090	0.75000	0.01868
+	1.31559	+.824	0.79523	0.77777	0.04523
+	1.44947	+.908	0.81823	0.80555	0.04045
+	1.49448	+.936	0.82558	0.83333	0.02003
+	1.56756	+.982	0.83711	0.86111	0.02399
+	1.89865	+1.190	0.88302	0.88888	0.02191
+	2.15469	+1.350	0.91160	0.91666	0.02272
+	2.48018	+1.554	0.94000	0.94444	0.02333
+	2.85253	+1.788	0.96312	0.97222	0.01868
+	3.39181	+2.126	0.98325	0.99999	0.01674

D = 0,927

5. CONCLUSÕES

O presente trabalho nos levou às seguintes principais conclusões:

- 1) Embora as variáveis independentes tenham proporcionado ajustes de modelos de regressões lineares múltiplas com alta explicabilidade para as perdas de peso (PPES), e parâmetros altamente significativos, concluímos que os mesmos não são adequados, em virtude das acentuadas tendências indicativas de não linearidade e variância não constantes, demonstradas pelas análises dos resíduos. Para estes tipos de dados, o ajustamento através de modelos não lineares ou de simulação poderão conduzir a estimativas mais realistas do verdadeiro comportamento das perdas de peso.

- 2) Mesmo tendo sido considerado não adequado, a utilização do método de todas as regressões possíveis apresentou informações importantes com respeito às variáveis que estão mais relacionadas com a PPES, as quais não poderiam ser obtidas através dos procedimentos usuais de STEPWISE.
- 3) A estatística C_p se mostrou realmente eficiente na seleção daquelas equações mais satisfatórias do ponto de vista de porcentagem de explicabilidade.
- 4) A determinação do grau de multicolinearidade, através do cálculo da estatística Ψ também foi considerada relevante, por informar que algumas das variáveis independentes são dispensáveis por estarem altamente correlacionadas com outras, e também considerando imprópria a utilização dos métodos STEPWISE.
- 5) As análises dos dados pelo método dos componentes principais, embora também consideradas inadequadas, serviram para confirmar as opiniões de BELSLEY *et alii* (1980), de que a multicolinearidade não pode ser detectada pelas análises dos resíduos.
- 6) As análises dos resíduos foram os instrumentos fundamentais que nos levaram a concluir como inadequada a utilização das análises de regressões lineares múltiplas e componentes principais nas avaliações das perdas de peso, confirmando realmente a importância de suas aplicações.

- 7) Alguns dos resultados para o teste dos chorrilhos consideraram não tendenciosos os modelos que pelas análises gráficas assim foram considerados, nos levando a concluir que seus resultados são medianamente satisfatórios.
- 8) Os resultados obtidos para normalidade dos resíduos pelas aplicações do teste de Lilliefors, foram satisfatórios e informativos.
- 9) A utilização dos programas desenvolvidos em linguagem BASIC para microcomputadores conduziu a bons resultados dos pontos de vista de capacidade, precisão e tempo de processamento.

6, BIBLIOGRAFIA

AFIFI, A.A. e S.P. AZEN, 1979. Statistical Analysis. Nova York, John Wiley, 442 pp.

ANSCOMBE, F.J., 1973. Graphs in Statistical Analysis. American Statistician, 27: 17-21.

ANSCOMBE, F.J. e J.W. TUKEY, 1963. The Examination and Analysis of Residuals. Technometrics, 5: 141-160.

BASKERVILLE, J.C. e J.H. TOOGOOD, 1982. Guided Regression Modeling for Prediction and Exploration of Structure with Many Explanatory Variables. Technometrics, 24: 9-18.

BELSLEY, D.A.; E. KUH e R.E. WELSCH, 1980. Regression Diagnostics. Nova York, John Wiley, 292 pp.

- CAMPOS, H., 1979. Estatística Experimental Não-Paramétrica. ESALQ/USP, 343 pp.
- CHATTERJEE, S. e B. PRICE, 1977. Regression Analysis by Example. Nova York, John Wiley, 228 pp.
- COOLEY, W.W. e P.R. LOHNES, 1962. Multivariate Procedures for the Behavioral Sciences. Nova York, John Wiley, 211 pp.
- DANIEL, C. e F.S. WOOD, 1980. Fitting Equations to Data. 2ª edição, Nova York, John Wiley, 458 pp.
- DRAPER, N.R. e H. SMITH, 1966. Applied Regression Analysis. Nova York, John Wiley, 407 pp.
- FIGUEIREDO, G.C., 1977. Teste Baseado no Número Total de Chorrilhos. Piracicaba, ESALQ/USP. Departamento de Matemática e Estatística, 10 pp. (Mimeografado).
- FURNIVAL, G.M., 1971. All Possible Regressions with Less Computation. Technometrics, 13: 403-408.
- FURNIVAL, G.M. e R.W.M. WILSON, 1974. Regressions by leaps and bounds. Technometrics, 16: 499-511.
- GARSDALE, M.J., 1965. The Best Sub-set in Multiple Regression Analysis. Applied Statistical, 14: 196-200.
- GOODNIGHT, J., 1979. A Tutorial on the SWEEP Operator. American Statistician, 33:

- GORMAN, J.W. e R.J. TOMAN, 1966. Selection of Variables for Fitting Equations to Data. Technometrics, 4: 27-51.
- HAWKINS, D.M., 1981. A New Test Multivariate Normality and Homocedasticity. Technometrics, 23: 105-110.
- HOFFMANN, R. e S. VIEIRA, 1977. Análise de Regressão. São Paulo, HUCITEC, Ed. da USP, 339 pp.
- HOCKING, R.R., 1976. The Analysis and Selection of Variables in Linear Regression. Biometrics, 32: 1-40.
- HOCKING, R.R. e R.N. LESLIE, 1967. Selection of the Best Sub-set in Regression Analysis. Technometrics, 2: 531-540.
- HUBER, P.J., 1981. Robust Statistics. Nova York, John Wiley, 308 pp.
- JOHNSON, N.L. e F.C. LEONE, 1977. Statistics e Experimental Design in Engineering and the Physical Sciences. 2ª edição, Nova York, John Wiley, 1082 pp.
- LEVENTHAL, L.A., 1979. 6502 Assembly Language Programming. Berkeley, Osborne, McGraw-Hill, 471 pp.
- LINDGREN, B.W., 1976. Statistical Theory. 3ª edição, Nova York, McMillan Publishing, 614 pp.
- MALLOWS, C.L., 1964. Choosing Variables in a Linear Regression: a Graphical Aid. Apresentado no Central Regional Meeting of the Institute of Mathematical Statistics, Manhattan, Kansas.

MALLOWS, C.L., 1973. Some Comments on C_p . Technometrics, 15: 661-675.

MARQUADT, D.W. e P.R. SNEE, 1975. Ridge Regression in Practice. American Statistician, 29: 3-20.

MATTIOLLI, J.C., 1978. Efeitos da intensidade da população de *Sitophilus oryzae* (Linné, 1763) (Coleoptera, curculionidae) e do período de armazenamento sobre diferentes variedades de milho. Curitiba, Universidade Federal do Paraná, 192 pp. (Dissertação de Mestrado).

MORRISON, D.F., 1967. Multivariate Statistical Methods. 2ª edição, Nova York, McGraw-Hill, 415 pp.

MOSTELLER, F. e J.W. TUKEY, 1977. Data Analysis and Regression. Addison-Wesley, Mass., 428 pp.

OSBORNE, A. e C. DONAHUE, 1980. PET/CBM Computer Guide, 2ª edição, Berkeley, Osborne/McGraw-Hill, 501 pp.

SCHATZOFF, M.; R. TSAO e S. FIENBERG, 1968. Efficient Calculation of All Possible Regressions. Technometrics, 10: 769-779.

SEBER, G.A.F., 1977. Linear Regression Analysis. Nova York, John Wiley, 465 pp.

SUNG, H.P., 1981. Collinearity and Optimal Restrictions on Regression Parameters for Estimating Responses. Technometrics, 23: 289-296.

WEISBERG, S., 1980. Applied Linear Regression. Nova York, John Wiley, 283 pp.

WEISBERG, S., 1981. A Statistic for Allocating C_p to Individual Cases. Technometrics, 23: 27-32.

YATES, F., 1937. Design and Analysis of Factorial Experiments. Harpenden, Graland: Imperial Bureau of Soil Science.*

* Trabalho não consultado no original.

7. APÊNDICE

APENDICE 1 - Programa STEP.COMP.

```

1 GOT0870
49 NU=9
50 FORI=1TONU:READV$(I):NEXT
57 CB$=" OB "+V$(1)+" "+V$(2)+"
60 RETURN
100 REM***CALCULO DA DISTRIBUICAO DOS CHORRILHOS ***
105 FI=1:FO=1:VX=HX-L:IFVX=0 OR L=0THENFI=1:GOTO 125
107 IFN<VX OR NX<LTHENFI=0:GOTO 125
110 FI=1:FO=FI:VX=NX-L:IFVX>LTHEN NE=L:L=VX:VX=NE
120 FORF=(L+1)TONX:FI=FI*F:NEXT:FORF=1TOVX:FO=FO*F:NEXT:FI=FI/FO
125 RETURN
290 PRINT"O POSICIONE O PAPEL NO PICOTE"
295 GETX$:IFX$=""THEN295
300 RETURN
870 PRINT"J"
880 POKES9468,14
890 DIMTT(73)
900 FORI=1TO57:READTT(I):NEXT
910 READP,B1,B2,B3,B4,B5
1000 REM ***STEPWISE ***
1010 REM ***EXPERIENCIA ***
1040 REM ***IDENTIFICACAO DA ANALISE ***
1050 PRINT"QUAL O TITULO DA ANALISE?":INPUTTZ$
1060 REM ***ESPECIFICACOES ***
1061 PRINT"UMERO DE VARIABEIS QUE VAO ENTRAR":INPUTM
1062 PRINT"UMERO DE VARIABEIS A SEREM UTILIZADAS NO MODELO":INPUTNU:PRINT
1064 PRINT"UMERO DE OBSERVACOES":INPUTNO:PRINT
1065 PRINT"UMERO DE TRANSFORMACOES A SEREM FEITAS":INPUTNT:PRINT
1066 PRINT"UMERO DE CONSTANTES UTILIZADAS":INPUTNC:PRINT
1067 PRINT">ALOR DE - PARA INCLUSAO DE VARIABEIS":INPUTFI:PRINT
1070 PRINT">ALOR DE - PARA EXCLUSAO DE VARIABEIS":INPUTFO:PRINT
1071 DIMDA(40,20),SI(40),B(NU),ID(NU),RI(NU+3,NU+3),XB(NU),SG(NU)
1072 DIMCO(12),IT(NV),JT(NV),LT(NV),KT(NV),CV(NV):V=.000000005
1080 IFI<FOTHEN5020
1090 IFNT<0THEN5000
1100 IFNT=0THEN1170
1110 REM ***ENTRADA DAS TRANSFORMACOES ***
1112 FORI=1TONT:PRINT"TRANSF I":PRINT"II":INPUTIT(I)
1114 PRINT"ABCDEFGHIJ":INPUTJT(I):PRINT"KLMNOPQRST":INPUTKT(I)
1120 PRINT"UVWXYZ":INPUTLT(I):NEXT:PRINT
1130 IFNC<0THEN5000
1140 IFNC=0THEN1170
1150 REM ***ENTRADA DAS CONSTANTES ***
1160 PRINT:PRINT:FORI=1TONC:PRINT"CONSTANTE I":INPUTCO(I):NEXT
1170 OB=NO:NI=NI-1:OPEN1,4:PRINT#1,"@ TZ$:PRINT#1:CLOSE1
1180 FORI=1TONU:XB(I)=0:FORJ=1TONU:RI(I,J)=0:NEXT:NEXT
1181 PRINT"LEITURA VIA TECLADO OU DISKETTE T/D?"
1182 GETX$:IFX$=""THEN1182
1183 IFX$<"T" AND X$<"D" THEN1182
1184 IFX$="D"THEN1191
1185 INPUT# DE VARIABEIS A ENTRAR":NW:PRINT"J":FORI=1TONO:PRINT"OBS. I";
1187 FORJ=1TONN:PRINTTAB(8)"X(J)":INPUTDA(I,J):NEXT:PRINT:NEXT:GOTO1205
1190 REM ***ENTRADA DAS OBSERVACOES ***
1191 PRINT:PRINT
1192 INPUT"QUAL ARQUIVO":AR$:INPUT"QUAL DRIVE":AS$:AR$=AS$+" "+AR$+",S,R"
1195 OPEN5,8,2,AR$
1196 FORI=1TONO
1200 FORJ=1TONV:INPUT#5,DA(I,J):NEXT:NEXT

```



```

1205 FORI=1TONU
1210 IFNT<0THEN5000
1220 IFNT=0THEN1400
1230 REM ***** TRANSFORMACAO DOS DADOS *****
1240 FORM=1TOHT:II=IT(M):JJ=JT(M):KK=KT(M):LL=LT(M)
1250 DHI1GOTO1260,1270,1280,1290,1300,1310,1320,1330,1340,1350,1360
1260 DA(I,JJ)=DA(I,KK):GOTO1390
1270 DA(I,JJ)=-DA(I,KK):GOTO1390
1280 DA(I,JJ)=LOG(DA(I,KK)):GOTO1390
1290 DA(I,JJ)=1/DA(I,KK):GOTO1390
1300 DA(I,JJ)=DA(I,KK)+DA(I,LL):GOTO1390
1310 DA(I,JJ)=DA(I,KK)*DA(I,LL):GOTO1390
1320 DA(I,JJ)=DA(I,KK)/DA(I,LL):GOTO1390
1330 DA(I,JJ)=DA(I,KK)+CO(LL):GOTO1390
1340 DA(I,JJ)=DA(I,KK)*CO(LL):GOTO1390
1350 DA(I,JJ)=DA(I,KK)*CO(LL):GOTO1390
1360 IFDA(I,KK)<100THEN1380
1370 DA(I,JJ)=90:GOTO1390
1380 Z=DA(I,KK):DA(I,JJ)=ATN(SQR(Z/(100-Z)))*180/3.141592654
1390 NEXT
1400 FORJ=1TONU:XB(J)=XB(J)+DA(I,J):FORK=1TONU
1470 RI(J,K)=RI(J,K)+DA(I,J)*DA(I,K):NEXT:NEXT
1480 NEXT:X$="I X":R$="I Y"
1490 FORI=1TONU:SG(I)=(RI(I,I)-Y(BD)*12/OB)*10.5:NEXT
1500 FORI=1TONU:FORJ=1TONU:RI(I,J)=(RI(I,J)-XB(I)*XB(J)/OB)/(SG(I)*SG(J))+V
1510 NEXT:NEXT
1515 PRINT"TESEJA A IMPRESSAO DOS DADOS S/T I":CLOSE1
1516 GETY$:IFY$=""THEN1516
1517 IFY$<>"S" AND Y$<>"I"THEN1516
1518 IFY$="I"THEN1600
1520 OPEN1,4:PRINT#1:PRINT#1
1530 PRINT#1:OPEN2,4,1:OPEN3,4,2:PRINT"-FORMATO F? ?"
1531 INPUTFK:A1=INT(FK)-1-INT((FK-INT(FK))*10+.05)
1532 FORI=1TOA1:ZY#=ZY#+9":NEXT
1533 ZY#=ZY#+".":FORI=1TOINT((FK-INT(FK))*10+.05):ZY#=ZY#+9":NEXT:ZY#=ZY#+". "
1534 W$="ZZZ 99.999 999.9 ":FORI=3TO4:W$=W$+ZY$:CB$=CB$+V$(I)+":NEXT
1535 PRINT#1,CB$
1560 PRINT#3,W$
1570 FORI=1TONU:PRINT#2,I:FORJ=1TONU
1580 PRINT#2,DA(I,J):NEXT:PRINT#1,CHR$(13):NEXT:PRINT#1:CLOSE2:CLOSE3
1600 REM ***** CALCULO DAS MEDIAS, E DESVIO *****
1605 STOP
1610 T$="-----"
1620 U$="-----"
1630 V$="-----"
1640 Y$="IVARIÁVEL 11      MÉDIA 11      DESVIO 11  PADRÃO 11  COEF. 11"
1645 Y$=Y$+"-----"
1650 FORI=1TONU:XB(I)=XB(I)/OB:SG(I)=SG(I)/(OB-1)*10.5:CV(I)=SG(I)*100/XB(I)
1660 NEXT:TS=SG(NU)*2*(OB-1)
1670 CLOSE1:OPEN1,4:PRINT#1,TAB(10)T$:PRINT#1,TAB(10)V$:PRINT#1,TAB(10)U$
1680 W$="AAAAAAA RI 99999.99999  RI 9999.99999  RI 99999.99  RI"
1690 OPEN2,4,1:OPEN3,4,2
1700 PRINT#3,TAB(10)W$
1705 I=1
1710 PRINT#2,"I  "V$(I),CHR$(29),XB(I),SG(I),CV(I):PRINT#1:I=I+1
1720 IFI=NUTHENPRINT#2,"I  "V$(NU),CHR$(29),XB(I),SG(I),CV(I):GO
1730 GOTO171
1740 PRINT#1:PRINT#1,CHR$(141)TAB(10)V$:CLOSE2
1750 PRINT#1:PRINT#1,CHR$(1)"          MATRIZ DE 1"
1760 PRINT#1:OPEN2,4,1:OPEN3,4,2
1764 W$="":FORI=1TONU:W$=W$+"S2.9999  "

1790 FORI=1TONU:
1800 PRINT#2,RI(I,J):NEXT:PRINT#1,CHR
1810 PRINT#2:CLOSE2:PRINT#3
1820 DF=OB-1:HF=-1:FORI=1TONU:SG(I)=SG(I)*(OB-1)*10.5:

```

```

1830 PRINT#1,CHR$(1)"          REGRESSAO PASSO AO PASSO"
1840 PRINT#1
1845 NF=NF+1
1850 SD=((RI(NU,NU)/DF)*10.5)*SG(NU):DF=DF-1:IFDF<=0THEN5010
1855 NN=0:VN=0:VX=0
1860 FORI=1TONI
1870 VI=RI(I,NU)*RI(NU,I)/RI(I,I)
1880 IFVI<0THEN1920
1890 IFVI=0THEN1960
1900 IFVI<=VWTHEN1960
1910 VX=VI:NN=I:GOTO1960
1920 NN=NN+1:ID(NN)=I
1930 REM ***** CALCULA OS COEFICIENTES B *****
1940 B(NN)=RI(I,NU)*SG(NU)/SG(I):SI(NN)=(SD*RI(I,I)*10.5)/SG(I):IFVN>0THEN5000
1945 IFVN=0THEN1950
1946 IFVI<=VWTHEN1950
1950 VN=VI:NM=I
1960 NEXT
1970 IFNN<0THEN5000
1980 IFNN=0THEN2400
1990 REM ***** CALCULA O TERMO CONSTANTE *****
2000 BS=XB(NU):FORI=1TONN:J=ID(I):BS=BS-B(I)*X(J):NEXT
2010 IFNE<0THEN5000
2020 IFNE=0THEN2420
2025 IFNN<0NU-1THEN2070
2030 OPEN4,4:CMD4:PRINTCHR$(1)"ENTRADA #\"NP;\" - ENTRA A VARIAVEL - X\"V$(K)
2040 PRINT\"RESVIO PADRAO RESIDUAL\"=\"SD:R=(1-RI(NU,NU))*10.5:R2=R*2
2050 PRINT\"COEFICIENTE DE CORRELACAO\"=\"R
2055 PRINT\"% DE VARIACAO EXPLICADA\"=\"R2*100
2060 PRINT\"TERMO CONSTANTE\"=\"BS:PRINT#4:CLOSE4
2070 S1=R2*TS:S2=TS-S1:G1=OB-2-DF:G2=DF+1:Q1=S1/G1:Q2=S2/G2:G3=OB-1:F=Q1/Q2
2075 IFNN<0NU-1THEN2300
2080 OPEN4,4:CMD4:PRINTCHR$(1)" ANÁLISE DA VARIANCA\":PRINT#4:CLOSE4
2090 F1$=\"REGRESSAO\":F2$=\"RESIDUO\":F3$=\"TOTAL\"
2100 F4$=
2110 F5$=\"C.VARIACAO   G.L.      S.Q.          Q.M.          F\"
2120 OPEN2,4,1:OPEN3,4,2:OPEN4,4:CMD4:PRINTF4$:PRINTF5$:PRINTF4$
2130 W$=\"AAAAAAAA 999 99999.9999          9999.99999 9999.99\"
2140 PRINT#3,W$
2150 PRINT#2,F1$,CHR$(29),G1,S1,Q1,F
2160 PRINT#2,F2$,CHR$(29),G2,S2,Q2;
2170 CMD4:PRINTF4$
2175 W$=\"AAAAAAAA 999 999999.9999\":PRINT#3,W$
2180 PRINT#2,F3$,CHR$(29),G3,TS:CMD4:PRINTF4$
2185 PRINT#2:CLOSE2:PRINT#3:CLOSE3:PRINT#4:CLOSE4

2190 U$=":
2200 V$=":
2210 Z$=":
2215 H$=":U$=U$+H$+"":V$=V$+H$+"":Z$=Z$+H$+"":
2220 T$=\" VARIAVEL COEFICIENTE B ( ERRO PADRAO ) INTERVALO DE\"
2230 T$=T$+\" CONFIANCA \"
2240 OPEN2,4,1:OPEN3,4,2:OPEN4,4
2250 CMD4:PRINTU$:PRINTT$:PRINTV$:
2260 FORI=1TONN:J=ID(I):IFG2>60THENTQ=2:GOTO2260
2270 TQ=TT(G2)
2290 W$=\"AAAAAAAA d1 9999.999999- Ad199999.99999 d1999.99999-AAA999A999.\"
2300 W$=W$+\"99999- d1\"
2310 TZ$=\"<=B\":H$=\"<=\"
2320 PRINT#3,W$
2330 I=1
2335 G=ID(I)
2336 LI=B(I)-SI(I)*TQ:LS=B(I)+SI(I)*TQ
2337 IFABS(B(I)/SI(I))>TRTHEND$=\"\":GOTO2340
2338 D$=CHR$(254)
2340 PRINT#2,\" I \"V$(G),CHR$(29),B(I),D$,CHR$(29),SI(I),LI,TZ$,CHR$(29),G,H$:
2345 PRINT#2,CHR$(29),LS;

```

```

2350 I=I+1:IFI=<N1THENCMD4:GOTO2335
2370 CMD4:PRINT#2:PRINT#2:CLOSE#2:PRINT#3:CLOSE#3:PRINT#4:CLOSE#4
2380 FV=VN*DF/RI(NU,NU):IF(FV+V)<=0THEN2400
2390 K=NM:NE=0:DF=DF+2:GOTO2440
2400 FV=VN*DF/RI(NU,NU)-VX:IFFV<=FITHEN2560
2410 K=NX:NE=K:GOTO2440
2420 OPEN#4:CMD4:PRINTCHR$(1)"ENTRADA # "NP;" - SNAI A VARIÁVEL - X"K
2430 GOTO2040
2440 FORI=1TONU:IFI=KTHEN2480
2450 FORJ=1TONU:IFJ=KTHEN2470
2460 RI(I,J)=RI(I,J)-RI(I,K)*RI(K,J)/RI(K,K)
2470 NEXT
2480 NEXT
2490 FORJ=1TONU:IFJ=KTHEN2510
2500 RI(K,J)=RI(K,J)/RI(K,K)
2510 NEXT
2520 FORI=1TONU:IFI=KTHEN2540
2530 RI(I,K)=-RI(I,K)/RI(K,K)
2540 NEXT
2550 RI(K,K)=1/RI(K,K):GOTO1845
2560 NEXT:FORL=1TO3
2570 DA(L,NU+1)=BS:FORI=1TONH:J=ID(I):DA(L,NU+1)=DA(L,NU+1)+E(I)*DA(L,J):NEXTI
2580 DA(L,NU+2)=DA(L,NU)-DA(L,NU+1)
2585 NEXTL
2587 FORL=1TONO:SI(L)=L:DA(L,NU+3)=DA(L,NU+2):NEXT
2590 PRINT"TRIEÑA EM RELACAO A QUAL VARIÁVEL":INPUTQ:Q=Q-NU
2660 M=NO-1
2670 L=0
2680 FORI=1TOM
2690 IFDA(I,NU+0)<=DA(I+1,NU+0)THEN2720
2700 FORJ=1TO(CHU+2):NE=DA(I,J):DA(I,J)=DA(I+1,J):DA(I+1,J)=NE
2710 NEXT
2711 L=L+1:VN=SI(I):SI(I)=SI(I+1):SI(I+1)=VN
2720 NEXT
2730 M=M-1:IFL=1THEN2670
2740 M=NO-1
2750 L=0
2760 FORI=1TOM
2770 IFDA(I,NU+3)<=DA(I+1,NU+3)THEN2800
2780 NE=DA(I,NU+3):DA(I,NU+3)=DA(I+1,NU+3):DA(I+1,NU+3)=NE
2790 L=L+1
2800 NEXT
- 2810 M=M-1:IFL=1THEN2750
2820 FORI=1TONO:T1=T1+DA(I,NU+1)*DA(I,NU+2):T2=T2+DA(I,NU+1)^2*DA(I,NU+2)
2830 T3=T3+DA(I,NU+1)*DA(I,NU+2)^2:NEXT
2835 CH=1
2840 FORI=1TO(NO-1):IF(DA(I,NU+2)*DA(I+1,NU+2))>0THEN2860
2850 CH=CH+1
2860 NEXT
2862 FORI=1TONO:IFDA(I,NU+2)<0THENH1=N1+1
2863 NEXT
2870 U$=" "
2875 I=I+1
2880 V$=" "
2890 Z$=" "
2895 E$="!"
) I"
2910 T$=T$+" RESIDUO " I"
2916 N$="N1 999 N199999.9999- 3199999.9999- 219999.99999 N1 "
2919 N$=N$+" N1 999 N199999.99999 N1
2920 OPEN#4:OPEN#2,4,1:OPEN#3,4,2
2930 PRINT#4," VALORES ORDENADOS EM RELACAO A Y. REST.O"
2940 PRINT#4,U$:PRINT#4,T$:PRINT#4,V$
2950 PRINT#3,N$
2960 I=1
2970 PRINT#2,SI(I),DA(I,NU),DA(I,NU+1),DA(I,NU+2)/SD,SI(I),DA(I,NU+2):I=I+1

```

```

2982 IF I<N0 THEN 2970
2990 PRINT#2,SI(I),DA(I,NU),DA(I,NU+1),DA(I,NU+2)/SD,SI(I),DA(I,NU+2);
2995 PRINT#4
3000 PRINT#4,Z#:PRINT#2:CLOSE2:PRINT#3:CLOSE3
3010 PRINT#4,CHR$(1)CHR$(1)"T"CHR$(129)CHR$(129)"11="T1
3020 PRINT#4,CHR$(1)CHR$(1)"T"CHR$(129)CHR$(129)"21="T2
3030 PRINT#4,CHR$(1)CHR$(1)"T"CHR$(129)CHR$(129)"12="T3
3040 PRINT"NUMERO DE CHORRILHOS="CH
3045 POKE59468,12
3050 PRINT#4:CLOSE4
3055 IF N0<=40 THEN K=3:N2=N0-N1:GOTO3090
3060 NT=2*N1*N2/N0+1:NC=2*N1*N2*(2*N1*N2-N1-N2)
3070 NC=NC/(N1+N2)12*(N1+N2-1):Z=(CH-NT+.5)/SQR(NC)
3080 IF Z<=-1.96 OR Z>=1.96 THEN K=2
3085 C#=""
3084 GOSUB4000
3085 FX=FX*100
3090 PRINT#1:CLOSE1:GOSUB290:OPEN1,4:CMD1
3092 PRINTCHR$(1)"***** TEXTE DOS [C]HORRILHOS[ ] *****":PRINT:PRINT
3093 PRINTCHR$(1)"[R] ="CH
3100 PRINTCHR$(1)"N"CHR$(129)"1";:PRINTCHR$(1)"="N1
3102 PRINTCHR$(1)"N"CHR$(129)"2";:PRINTCHR$(1)"="N2
3107 IF N0<=40 THEN 3160
3110 PRINTCHR$(1)"Z ="Z
3115 FX=INT(100*FX)/100
3120 PRINTCHR$(1)"[Z]=[Z]"="FX"%"
3130 PRINTCHR$(1)"P(Z<=[Z])="100-FX"%"
3135 PRINTCHR$(1)"H&N=0MS ERROS NAO SAO INDEPENDENTES[ ]"
3137 PRINTCHR$(1)"H&N=1MS ERROS SAO INDEPENDENTES[ ]"
3140 IF K=2 THEN PRINTCHR$(1)"ACEITAMOS [H&N ]"
3150 PRINTCHR$(1)"ACEITAMOS [H&N ]"
3155 GOTO3300
3160 NX=N0:L=N1:GOSUB100:S2=FI
3170 IF(N0/2-INT(N0/2))>0 THEN TQ=N0-1:SD=N0:GOTO3180
3175 TQ=N0:SD=N0-1
3180 I=0
3182 I=I+2:IFI>TQ THEN 3190
3183 GOSUB3090:TT(I)=S1:GOTO3182
3190 I=1
3192 I=I+2:IFI>SD THEN 3196
3195 GOSUB3090:TT(I)=S1:GOTO3192
3196 Y#=""
3197 Z#=""
3198 T#=""
3199 U#="" | [R] | P([R]=R.[ ] | P([R]<=R.[ ] | P([R]>=R.[ ] | "
3200 OPEN2,4,1:OPEN3,4,2:W#=""@I ZZ @I Z.99999 @I Z.99999@ @I Z.99999@ @
3210 PX=0:PY=1:A#=CHR$(254):B#=""
3220 CMD1:PRINT:PRINTCHR$(1)" DISTRIBUICAO DOS [C]HORRILHOS[ ]"
3225 PRINT:PRINT:PRINTY#:PRINTU#:PRINTZ#
3230 PRINT#3,W#
3235 I=2:PY=1:TT(1)=0
3240 PX=PX+TT(I):PY=-TT(I-1)+PY:IFI=CH THEN 3260
3250 PRINT#2,I,TT(I),PX,A#,CHR$(29),PY,A#,CHR$(29):I=I+1:IFI=NO THEN 3270
3257 A#=0:GOTO3240
3260 PRINT#2,I,TT(I),PX,B#,CHR$(29),PY,B#,CHR$(29):I=I+1:IFI=NO THEN 3270
3265 A#=0:GOTO3240
3270 A#=0:IF N0=CH THEN 3275
3274 B#=CHR$(254)
3275 PX=PX+TT(I):PY=-TT(I-1)+PY
3280 PRINT#2,I,TT(I),PX,B#,CHR$(29),PY,B#,CHR$(29):CMD1:PRINTT#
3282 PRINT#1:CLOSE1
3285 PRINT#2:CLOSE2:PRINT#3:CLOSE3
3290 GOSUB290:OPEN1,4:CMD1
3300 PRINT:PRINTCHR$(1)" TEXTE DE [L]MILLIEFORS[ ] - [N]ORMALIDADE[ ]"
3310 HE=0:NX=0:FOR I=1 TO N0:NX=NX+DA(I,NU+3)12:
NEXT 3320 SD=SQR(Q2)

```

```

3330 Y#="
3340 U#="
3350 Z#="
3360 W#="M 99999.99999 M 99.999 M 2.99999 M 2.99999 M 2.99999
3370 T#="| RESIDUO | Z | F(X) S(X) |F(X)-S(X)|
3373 K=-99999:PY=0
3375 OPEN2,4,1:OPEN3,4,2
3376 CMD1:PRINTY#:PRINTT#:PRINTU#
3377 PRINT#3,W#
3378 I=0:PY=0
3379 I=I+1
3380 Z=DA(I,HU+3)/SD:PY=PY+1/HO:GOSUB4000:FI=ABS(PX-PY)
3385 IFFI<ABS(PX-PY+1/HO)THENFI=ABS(PX-PY+1/HO)
3387 A=0:IFI=HO THEN3410
3390 PRINT#2,DA(I,HU+3),Z,PX,PY,FI:IFFI>KTHENK=FI
3400 A=0:GOTO3379
3410 PRINT#2,DA(I,HU+3),Z,PX,PY,FI:CMD1:PRINTZ#
3415 K=INT(K*10000)/10000:K#=STR$(K):K#="0"+MID$(K#,2,5)
3420 FORI=1TO70:READTT(I):NEXT
3422 T#=STR$(TT(HO)):T#="P(D)=0"+MID$(T#,2,4)+">=5%"
3425 CMD1:PRINTCHR$(1)T#
3430 IFK<TT(HO)THENPRINTCHR$(1)"D.="K#:GOTO3440
3435 PRINTCHR$(1)"D.="K#
3437 PRINTCHR$(1)"OXS RESIDUOS NAO APRESENTAM D":GOTO3345
3440 PRINTCHR$(1)"OXS RESIDUOS APRESENTAM "
3445 PRINTCHR$(1)"DISTRIBUICAO NORMALD"
3450 PRINT#1:CLOSE1:GOSUB290
3500 END

3630 REM ***** CALCULO DA PROBABILIDADE DE VALORES DE CHORRILHOS PAR *****
3640 L=I/2-1:NX=N1-1:GOSUB 100
3645 S1=2*FI:NX=N2-1:L=I/2-1:GOSUB 100
3650 S1=S1*FI/S2:RETURN
3660 REM ***** CALCULO DA PROBABILIDADE DE VALORES DE CHORRILHOS IMPAR *****
3670 NX=N1-1:L=(I-1)/2:GOSUB 100
3675 G1=FI:NX=N2-1:L=(I-3)/2:GOSUB 100
3680 G1=G1*FI:NX=N1-1:L=(I-3)/2:GOSUB 100
3685 G2=FI:NX=N2-1:L=(I-1)/2:GOSUB 100
3690 G2=G2*FI:G1=(G1+G2)/S2:RETURN
4000 LK=0
4002 IFZ=0THENPX=.5:RETURN
4005 IFZ<0THENLK=1
4007 T1=ABS(Z)
4010 Q=SOR(1/(2*.3.1415926)):T=1/(1+P#Z1)
4020 PX=Q*EXP((-Z1 T2)/2):PX=PX*T*(B1+T*(B2+T*(B3+T*(B4+T*B5))))
4030 IFLK=1THENRETURN
4040 PX=-PX+1:RETURN
4510 DATA12.706,4.302,3.182,2.776,2.570,2.446,2.354,2.306,2.262,2.228,2.201
4520 DATA2.178,2.160,2.144,2.131,2.119,2.109,2.100,2.093,2.086,2.079,2.073
4530 DATA2.068,2.063,2.059,2.055,2.051,2.048,2.045,2.042,2.039,2.036,2.034
4540 DATA2.032,2.03,2.029,2.026,2.024,2.022,2.021,2.019,2.018,2.016,2.015,2.014
4550 DATA2.012,2.011,2.01,2.009,2.008,2.007,2.006,2.005,2.004,2.004,2.003,2.002
4560 DATA.2316419,.3193815,-.3565638,1.781478,-1.821256,1.330274

4565 DATANINS,TARM,CARB,UMID,MASE,PROT,OLEO,IACD,PFES
4570 DATA.975,.842,.708,.624,.563,.519,.483,.454,.438,.409,.391,.375,.361,.349
4580 DATA.338,.327,.318,.309,.301,.294,.287,.281,.275,.269,.264,.259,.254,.25
4590 DATA.246,.242,.238,.234,.231,.227,.224,.221,.216,.215,.213,.210,.208,.206
4600 DATA.203,.200,.198,.196,.194,.192,.190,.188,.186,.185,.183,.181,.180,.178
4610 DATA.177,.175,.174,.172,.171,.170,.168,.167,.166,.164,.163,.162,.161,.160
5000 CLOSE4:OPEN4,8,15
5010 INPUT#4,A#,B#,C#,D#
5020 PRINTA#,B#,C#,D#
9999 END

```



```

1180 PRINT#1,CHR$(1)"    ** GOODNESS OF FIT+ C&P&O **":PRINT#1,CHR$(1)W#
1185 PRINT#1,CHR$(1)W#:PRINT#1:PRINT#1,"X"TZ#:PRINT#1,WL#
1190 REM *****ENTRADA DAS OBSERVAÇÕES *****
1195 PRINT"QUAL ARQUIVO DESEJA LER"
1196 INPUTAR#:INPUT"TRIVE #";AS#
1197 AR#=AS#+": "+AR#+",S,R":OPEN5,8,2,AR#
1200 FORI=1TOHD:FORJ=1TONV:INPUT#5,DA(I,J):NEXTJ,I:CLOSES:TI#="000000"
1205 FORI=1TOHD
1220 IFHT=0THEN1400
1230 REM *****TRANSFORMAÇÕES DOS DADOS *****
1240 FORM=1TOHT:II=IT(M):JJ=JT(M):KK=KT(M):LL=LT(M)
1250 ONII GOTO1260,1270,1280,1290,1300,1310,1320,1330,1340,1350,1360
1260 DA(I,JJ)=DA(I,KK):GOTO1390
1270 DA(I,JJ)=-DA(I,KK):GOTO1390
1280 DA(I,JJ)=LOG(DA(I,KK)):GOTO1390
1290 DA(I,JJ)=1/DA(I,KK):GOTO1390
1300 DA(I,JJ)=DA(I,KK)+DA(I,LL):GOTO1390
1310 DA(I,JJ)=DA(I,KK)*DA(I,LL):GOTO1390
1320 DA(I,JJ)=DA(I,KK)/DA(I,LL):GOTO1390
1330 DA(I,JJ)=DA(I,KK)+CO(LL):GOTO1390
1340 DA(I,JJ)=DA(I,KK)*CO(LL):GOTO1390
1350 DA(I,JJ)=DA(I,KK)+CO(LL):GOTO1390
1360 IFDA(I,KK)<100THEN1380
1370 DA(I,JJ)=90:GOTO1390
1380 Z=DA(I,KK):DA(I,JJ)=ATH(SQR(Z/(100-Z)))*180/3.141592654
1390 NEXT
1400 FORJ=1TONU:XB(J)=XB(J)+DA(I,J):FORK=1TONU:RI(J,K)=RI(J,K)+DA(I,J)*DA(I,K)
1470 NEXT:NEXT.
1480 NEXT:X#=" | X":R#=" | Y"
1490 FORI=1TONU:SG(I)=(RI(I,I)-XB(I)^2/OB)/OB*10.5:NEXT
1500 FORI=1TONU:FORJ=1TONU:RI(I,J)=(RI(I,J)-XB(I)*XB(J)/OB)/(SG(I)*SG(J))+Y
1510 NEXTJ,I
1600 REM ***** CALCULO DAS MEDIAS E DESVIOS PADROES *****
1610 T#=""
1620 U#=""
1630 V#=""
1640 Y#=""
1645 Y#=Y#+":VARR. J |"
1650 FORI=1TONU:XB(I)=XB(I)/OB:SG(I)=SG(I)/(OB-1)*10.5:CV(I)=SG(I)*100/XB(I)
1660 NEXT:TS=SG(U)*12*(OB-1)
1670 W#=""
1680 PRINT#1,T#:PRINT#1,Y#:PRINT#1,U#
1690 OPEN2,4,1:OPEN3,4,2
1700 PRINT#3,W#
1705 FORI=1TONU-1
1710 PRINT#2," | "CHR$(64+I),CHR$(29),XB(I),SG(I),CV(I):PRINT#1:NEXT
1720 PRINT#2," | Y",CHR$(29),XB(I),SG(I),CV(I):PRINT#1
1740 PRINT#1,V#
1745 HD=1
1750 T#=""
1755 Z#=" | VARIÁVEIS DO P&O C&P&O VARIÁVEIS DO P&O C&P&O |"
1757 Z#=Z#+": VARIÁVEIS DO P&O C&P&O |"
1760 U#=""
1765 Y#=""
1770 T#=T#+":U#=U#+":Y#=Y#+":
1775 GOSUB300
1800 REM ***** CALCULO DOS VALORES DOS Q.M. DOS 2ºNI MODELOS *****
1810 W#="AAAAAAAAAAAAA199119999.99A"
1820 W#=W#+":AAAAAAAAAAAAA199119999.99AA"
1830 W#=W#+":AAAAAAAAAAAAA199119999.99A1"
1835 L1=1:EE=2:PRINT#3,W#
1836 CB$(1)=" INDEHNHUAO":CB$(2)=" |":CB$(3)=" |"
1839 GOSUB50:R#=""
1840 FORK=1TONI
1845 CK(K)=2*(K-1)-1:LK(K)=2*(K+CK(K)):B#=B#+":0"
1850 GOSUB60:NEXTK

```

```

1860 DM=RA(NU, NU)*TS:DN=DM/(NO-NU)
1861 MP(1)=1:CP(1)=TS/DN-NO+2:
1865 GOSUB58:Z=1
1870 FORU=1TO2#HI-1
1880 FORK=1TONI
1890 IFUCX(K) THEN GOSUB59:GOTO1910
1900 A#=A#+ "0"
1910 IFCX(K)<LX(K) THEN 1922
1920 LX(K)=LX(K)+21(K+1):CX(K)=CX(K)+21K
1922 IX=VAL(MID$(A#, K, 1))
1923 JX=VAL(MID$(B#, K, 1))
1924 IF IX=JX THEN 1930
1925 IF IX>JX THEN Z=Z+1:GOTO1928
1926 Z=Z-1
1928 GOSUB60
1930 NEXTK
1935 B#=A#:A#="" :MP(EE)=Z
1940 CP(EE)=RA(NU, NU)*TS/DN-(NO-2*MP(EE))+.00005
1941 IFCP(EE)<=MP(EE)+.00005 THEN CB$(EE)=CB$(EE)+"←"
1948 EE=EE+1:IFEE>3 THEN GOSUB85
1950 NEXTU:GOSUB95
1960 PRINT#1,"N°PROCESSAMENTO "TI#" I.\.
* 2010 PRINT#1:CLOSE1:CLOSE2:CLOSE3
4999 END

```


APÉNDICE 3 - Programa AUTO-JACOBI.

```

1 GOTO1000
40 DATA1,33,72,60,3,1
42 FORH=1TO6:READA:CE#=CE#+CHR#(A):HEXT
43 OFENS,4,5:PRINT#5,CE#
44 RETURN
50 MM=N-1
52 KY=0:FORJ=1TOMM:IFA(I,I)>=A(I+1,I+1)THEN#57
54 MM=A(I,I):A(I,I)=A(I+1,I+1):A(I+1,I+1)=MM:KY=1
55 FORJ=1TON:MM=A(I,J):AA(I,J)=AA(I+1,J):AA(I+1,J)=MM:NEXT
57 NEXT:MM=MM-1:IFKY=1THEN#52
58 RETURN
100 REM ***** CALCULA AUTOVALORES *****
110 EPS=1E-20
120 IFN<=1THENRETURN
130 XN=0
140 FORI=1TON
150 FORJ=1TON
160 C(I,J)=A(I,J)
170 IFI=JTHEN#190
180 XN=XN+ABS(A(I,J))
190 NEXTJ,I
200 XF=XN*1.E-24
210 NR=0
220 FORI=1TON
230 FORJ=1TON
240 IFI<>JTHENAA(I,J)=0:GOTO260
250 AA(I,J)=1
260 NEXTJ,I
270 MAX=0
280 FORI=1TON-1
290 FORJ=I+1TON
300 IFMAX>ABS(A(I,J))THEN#320
310 MAX=ABS(A(I,J)):NI=I:NJ=J
320 NEXTJ,I
330 IFMAX<=EPSTHENRETURN
340 IFABS(A(NI,NI))-A(NJ,NJ)>>XPTHEN#380
350 Y=0:X=A(NI,NJ)/ABS(A(NI,NJ))
360 XS=X*SQR(0.5):XC=SQR(0.5)
370 GOTO450
380 ALFA=2*A(NI,NJ)/(A(NI,NI)-A(NJ,NJ))
390 RR=SQR(1+ALFA*ALFA)
400 SI=ALFA/ABS(ALFA)
410 X=ALFA/RR
420 Y=1/RR
430 XC=SQR(0.5*(1+Y))
440 XS=SI*SQR(0.5*(1-Y))
450 AUX=XC*XC
460 NR=NR+1:PRINTNR
470 BUX=XC*XS
480 CUX=XS*XS
490 B(NI,NI)=A(NI,NI)*AUX+2*A(NI,NJ)*BUX+A(NJ,NJ)*CUX
500 B(NJ,NJ)=A(NI,NI)*CUX-2*A(NI,NJ)*BUX+A(NJ,NJ)*AUX
510 B(NI,NJ)=A(NI,NJ)*Y+(A(NJ,NJ)-A(NI,NI))*X/2
520 A(NI,NI)=B(NI,NI)
530 A(NJ,NJ)=B(NJ,NJ)
540 A(NI,NJ)=B(NI,NJ)
550 A(NJ,NI)=B(NI,NJ)
560 FORK=1TON
570 IFK=NIORK=NJTHEN#650

```

```

590 B(NI,K)=A(NI,K)*YC+A(NJ,K)*XS
600 B(NJ,K)=-A(NI,K)*XS+A(NJ,K)*YC
610 A(NI,K)=B(NI,K)
620 A(NJ,K)=B(NJ,K)
630 A(K,NI)=A(NI,K)
640 A(K,NJ)=A(NJ,K)
650 NEXTK
660 FORK=1TON
670 BB(NI,K)=AA(NI,K)*XC+AA(NJ,K)*XS
680 BB(NJ,K)=-AA(NI,K)*XS+AA(NJ,K)*XC
690 AA(NI,K)=BB(NI,K)
700 AA(NJ,K)=BB(NJ,K)
710 NEXTK
720 GOTO270
1000 INPUT"ORDEN DA MATRIZ";N
1005 INPUT"# DE OBSERVAÇÕES";NO
1010 DIM A(N,N),AA(N,N),B(N,N),BB(N,N),C(N,N),X(NO,N+1),Z(NO,N+1)
1020 PRINT"*****LEITURA POR TECLADO OU DISKETE T/D?"
1030 GETX$:IFX$=""THEN1030
1040 IFX$<"T"ANDX$<"D"THEN1030
1050 IFX$="D"THEN1030
1060 FORI=1TON:IX(I)=I:FORJ=1TON
1070 PRINT"LINHA "I" COLUNA"J":INPUTA(I,J):NEXT:PRINT:NEXT:GOTO1100
1080 INPUT"ARQUIVO";AR$:OPEN2,S,2,"0:"+AR$+",S,R"
1090 FORI=1TON:FORJ=1TON:INPUT#2,A(I,J):NEXT:NEXT:CLOSE2
1100 PRINT"*****DESEJA CORRIGIR ALGUM DADO S/N?"
1110 GETX$:IFX$=""THEN1110
1120 IFX$<"S"ANDX$<"N"THEN1110
1130 IFX$="N"THEN1160
1140 INPUT"LINHA #";I:INPUT"COLUNA #";J:PRINT"VALOR HULO="A(I,J)
1150 INPUT"VALOR CORRETO";A(I,J):GOTO1100
1160 GOSUB100
1165 GOSUB50
1170 W$="":FORI=1TON:W$=W$+"S99.99999 " :NEXT
1175 GOSUB40
1180 OPEN1,4:OPEN2,4,1:OPEN3,4,2:PRINT#3,"A99A "+W$
1185 CH$=CHR$(254)+"#I"
1190 PRINT#1,"AUTO VALORES"
1200 FORI=1TON:FORJ=1TO1:PRINT#2,CH$,I,"=#I",A(I,I):NEXT:PRINT#1:NEXT
1205 PRINT#3,N$
1210 PRINT#1,"AUTO VETORES NAS LINHAS"
1220 FORI=1TON:FORJ=1TON:PRINT#2,AA(I,J):NEXT:PRINT#1:NEXT
1230 PRINT#1,"NUMERO DE ROTACOES="NR
1240 INPUT"ARQUIVO PADRONIZADO";AP$
1250 OPEN7,S,2,"0:"+AP$+",S,R"
1260 FORI=1TONO:FORJ=1TON+1:INPUT#7,X(I,J):NEXT:NEXT:CLOSE7
1270 FORJ=1TON:FORI=1TONO:FORK=1TON
1280 Z(I,J)=Z(I,J)+AA(J,K)*X(I,K)
1290 NEXT:NEXT:NEXT
1300 FORI=1TONO:Z(I,N+1)=X(I,N+1):NEXT
1310 PRINT#1:PRINT#1,"MATRIZ ORTOGONAL "
1320 FORI=1TONO:FORJ=1TON:PRINT#2,Z(I,J):NEXT:PRINT#1:NEXT
1440 CLOSE1:CLOSE2:CLOSE3

```

APENDICE 4 - Programa PLOT.RES.

```

100 REM ***** PROGRAMA PARA PLOTAR RESIDUOS *****
110 POKE59468,14:DIM X(100,15),Y(100)
120 INPUT"QUAL ARQUIVO A SER LIDO";AR$
130 INPUT"TRIVE #";D#
135 INPUT"# DE LINHAS, # DE COLUNAS";NL,NC
140 OPEN2,S,2,D#+" "+AR#+"",S,R
150 FORI=1TONL
160 FORJ=1TONC
170 INPUT#2,X(I,J)
180 NEXTJ,I
190 CLOSE2
200 INPUT"QUAL COLUNA E' A ASCISSA";AB
210 INPUT"QUAL COLUNA E' A ORDENADA";OD
220 PRINT"                AGUARDE UM INSTANTE !"
225 M=NL-1:AI=1E20:AS=-1E20
230 K=0:FORI=1TONL
231 IFX(I,AB)<=AI THENAI=X(I,AB)
232 IFX(I,AB)>=ASTHENAS=X(I,AB)
240 IFX(I,OD)>=X(I+1,OD) THEN310
250 FORJ=1TONC
260 TR=X(I,J)
270 X(I,J)=X(I+1,J)
280 X(I+1,J)=TR
290 NEXT
300 K=1
310 NEXT
320 M=M-1:IFK=1THEN230
330 PRINT"ORDENADA \MINIMA="X(NL,OD)
340 PRINT"ORDENADA \MAXIMA="X(1,OD)
350 PRINT"ASCISSA \MINIMA="AI
360 PRINT"ASCISSA \MAXIMA="AS
370 PRINT"-----IXE A ESCALA HORIZONTAL"
380 INPUT"LIMITE INFERIOR";LI
390 INPUT"LIMITE SUPERIOR";LS
400 PRINT"-----IXE A ESCALA VERTICAL"
410 INPUT"LIMITE INFERIOR";VI
420 INPUT"LIMITE SUPERIOR";VS
430 LH=30:LV=50
450 X#(0)="+":FORI=1TO10:X#(0)=X#(0)+"----+":NEXT
460 OPEN1,4,1:OPEN2,4,2:PRINT#2,"S9.99":PRINT#3
470 FORI=1TOLH:X#(1)="I":FORJ=1TOLV:X#(1)=X#(1)+" ":NEXT:NEXT
490 FORI=5TOLHSTEP5:X#(1)=" "+MID$(X#(1),2):NEXT
495 FORI=1TONL:K=1
500 ILX=INT((X(I,OD)-VI)*LH/(VS-VI)+.5)
505 IFILX>LVTHENILX=LV
510 COX=INT((X(I,AB)-LI)*LV/(LS-LI)+.5)+1
520 FX=VAL(MID$(X#(ILX),COX,1))
525 IFFX=KTHENK=K+1:GOTO525
530 X#(ILX)=MID$(X#(ILX),1,COX-1)+MID$(STR$(K),2)+MID$(X#(ILX),COX+1)
535 NEXTI:PRINT#2,0#
540 II=5:FORI=LHTO0STEP-1
570 PRINT#3,TAB(18)X#(I)
590 NEXT
600 PRINT"-----AVANCE O PAPEL PARA COMENTARIOS [A]"
610 GETY$:IFY#="" THEN610
620 IFY#<>"A" THEN610
640 INPUT"QUAL OME DA VARIAVEL";NV$
650 INPUT"COMENTARIO";C#
655 PRINT#3,"A" C#
660 PRINT#3,"ASCISSA="NV$ - "LI"A"LS" I"
680 CLOSE1:CLOSE2:CLOSE3

```