

ANÁLISE DA ADAPTABILIDADE ATRAVÉS DE REGRESSÃO LINEAR SEGMENTADA

1. FUNDAMENTOS¹

JOÃO GILBERTO CORRÊA DA SILVA²

RESUMO - O método de regressão linear simples de Finlay & Wilkinson (1963), estendido por Eberhart & Russell (1966), tem sido o mais utilizado para caracterização da estabilidade fenotípica e adaptabilidade de plantas cultivadas. Este método, entretanto, não permite a identificação dos distintos comportamentos das respostas de genótipos à variação de ambiente. Em particular, ele não é apropriado para identificar genótipos com as características desejáveis, isto é, que sejam responsivos a ambientes favoráveis ou melhorados e mantenham produção razoável em ambientes adversos. Com esse argumento, Silva & Barreto (1985) propõem representar a resposta de um genótipo à gama de ambientes por um gráfico composto de dois segmentos de reta, conectados no ponto correspondente ao índice de ambiente nulo. Este artigo expõe os fundamentos deste método e sua abordagem como uma extensão do método de regressão linear simples para proporcionar mais ampla flexibilidade para a caracterização dos distintos comportamentos das respostas de genótipos à variação de ambiente. A aplicação do método será ilustrada em artigo que segue.

Termos para indexação: estabilidade fenotípica, adaptabilidade ao ambiente, interação genótipo x ambiente.

ANALYSIS OF ADAPTATION THROUGH SEGMENTED LINEAR REGRESSION

1. FOUNDATIONS

ABSTRACT - The method of simple linear regression of Finlay & Wilkinson (1963), extended by Eberhart & Russell (1966), has been the most used to characterize the phenotypic stability and adaptability of cultivated plants. This method, however, does not allow the identification of the distinct behaviors of the responses of genotypes to the variation of the environment. Particularly, it is not appropriate to identify genotypes with the desirable characteristics, that is, which are responsive to favorable or improved environments and maintain reasonable production in adverse environments. With this reasoning, Silva & Barreto (1985) propose to represent the response of a genotype to a range of environments by a graph composed of two segments of straight line, connected in the point corresponding to the null environment index. This paper explains the foundations of this method and its approach as an extension of the simple linear regression method to allow broader flexibility for the characterization of the distinct behaviors of genotype responses to the environment variation. The application of the method will be illustrated in paper to follow.

Index terms: phenotypic stability, environmental adaptation, genotype x environment interaction.

INTRODUÇÃO

O objetivo de programas de melhoramento genético de plantas é, geralmente, a obtenção de genótipos que mantenham rendimentos consisten-

temente elevados na gama de ambientes que ocorrem nos locais e estações de cultivo de uma região. Uma característica comum dessas pesquisas é a alteração dos desempenhos relativos dos genótipos com variações de ambiente, ou seja, a presença de interação genótipo x ambiente. Como decorrência, essas pesquisas requerem elevados recursos, têm progresso lento e são de difícil avaliação.

O método de regressão linear tem sido o mais usado para a análise da interação genótipo x ambiente com vistas ao estudo da estabilidade fenotípica

¹ Aceito para publicação em 11 de janeiro de 1995.

² Eng. Agr., Ph.D., Prof. Titular (aposentado) Dep. de Matemática, Estatística e Computação, Inst. de Física e Matemática, Univ. Fed. de Pelotas, Caixa Postal 354, CEP 96010-900 Pelotas, RS. Bolsista do CNPq.

e adaptabilidade de genótipos. Essa abordagem, originalmente utilizada por Yates & Cochran (1938) na análise de grupos de experimentos, foi posteriormente desenvolvida por Finlay & Wilkinson (1963) e Eberhart & Russell (1966), e estendida por outros (Perkins & Jinks, 1968a, 1968b; Freeman & Perkins, 1971 e Shukla, 1972, por exemplo). Amplas revisões dos métodos de análise de adaptabilidade e estabilidade fenotípica são feitas por Easton & Clements (1973), Freeman (1973), Hill (1975), Westcott (1986) e Lin et al. (1986).

Basicamente, esse método de regressão linear adota o desempenho médio dos genótipos em um ambiente, ou o desvio deste em relação à média global de todos os ambientes, como um índice da produtividade do ambiente, e, então, representa a relação entre a resposta de um genótipo e o ambiente através de uma equação de regressão linear simples em relação a esse "índice de ambiente".

O índice de ambiente é usado como medida da produtividade do ambiente para contornar a complexidade da formulação de relações com as muitas variáveis ambientais que influenciam o rendimento, como precipitação pluvial, temperatura, fertilidade do solo, etc., de difícil mensuração e sobre as quais usualmente não são disponíveis dados. Como consequência, o coeficiente de regressão médio dos genótipos em consideração é igual a um, e a reta de regressão para as médias dos ambientes tem ajuste perfeito.

Finlay & Wilkinson (1963) caracterizam a estabilidade fenotípica e a adaptabilidade através de dois parâmetros: rendimento médio e coeficiente de regressão. Genótipos com coeficientes de regressão próximos de 1,0 têm *estabilidade média* em todos os ambientes. Quando também lhes correspondem rendimentos médios acima da média, têm *adaptabilidade geral*; se lhes são associados rendimentos médios abaixo da média são considerados *pobremente adaptados a todos os ambientes*. Genótipos com coeficientes de regressão maiores do que 1,0 têm *estabilidade abaixo da média*, e são *especificamente adaptados a ambientes favoráveis*; com coeficientes de regressão menores que 1,0 têm *estabilidade acima da média* e são *especificamente adaptados a ambientes desfavoráveis*.

Eberhart & Russell (1966) também consideram o coeficiente de regressão linear como uma medida de estabilidade fenotípica e propõem a variância do desvio da regressão como uma segunda medida. Eles caracterizam a adaptabilidade e a estabilidade fenotípica através de três parâmetros: rendimento médio, coeficiente de regressão e variância do desvio da regressão. Definem um genótipo como *estável*, se lhe correspondem coeficiente de regressão igual a um e variância do desvio da regressão zero, e como *desejável*, se é estável e de elevado rendimento médio.

A etapa preliminar desse método é a análise da variação conjunta dos ambientes. Se a interação genótipo x ambiente revela-se significativa, procede-se a sua análise com vistas a isolar os componentes sistemáticos atribuíveis às relações lineares entre as respostas dos genótipos e o índice de ambiente. A soma de quadrados da interação genótipo x ambiente é decomposta em dois componentes ortogonais: um, atribuível às diferenças entre as linhas de regressão ajustadas, e o outro, exprimindo os desvios acumulados dos valores observados em relação a essas linhas de regressão.

A significância da interação implica que pelo menos um desses dois componentes seja significativo. Se apenas o componente devido à heterogeneidade das linhas de regressão é significativo, conclui-se que a magnitude da interação é uma função linear do índice de ambiente para cada genótipo e que cada genótipo tem sua própria resposta linear característica à variação ambiental. Se, ao contrário, apenas o desvio é significativo, então, ou não existe uma relação linear simples, ou não existe qualquer relação entre a interação e a produtividade do ambiente expressa pelo índice de ambiente. Muito frequentemente, ambos os componentes mostram-se significativos. Nesse caso, as previsões referentes à interação genótipo x ambiente baseadas nas regressões lineares simples ainda poderão ter considerável valor prático (Perkins & Jinks, 1968a). Entretanto, o modelo linear não se terá revelado inteiramente satisfatório, já que uma fração significativa da interação genótipo x ambiente permanece inexplicada. A razão biológica para essa falha deve ser buscada pela apropriada manipulação dos dados (Hill, 1975).

Essa técnica de regressão tem mostrado utilidade em uma amplitude remarcável de casos. Suas principais vantagens derivam de sua simplicidade e habilidade para reduzir complexas interações a um conjunto ordenado de respostas lineares. Entretanto, ela tem sido objeto de críticas. Algumas delas referem-se a violações de pressuposições da metodologia estatística empregada, que não têm conseqüências sérias se são levadas em conta e devidamente consideradas as restrições que decorrem para as inferências derivadas (Baker, 1969; Freeman & Perkins, 1971; Hardwick & Wood, 1972; Freeman, 1973; Hill, 1975; Westcott 1986).

Hardwick & Wood (1972) salientam que, muitas vezes, desvios da regressão são relativamente elevados em decorrência de diferenças substanciais de respostas entre genótipos para mais de uma característica ambiental. De fato, parte do componente não explicado pelas relações lineares pode ser atribuído a características individuais dos genótipos; outra parte pode decorrer de extrema diversidade dos genótipos (Perkins & Jinks, 1968b). Um genótipo pode revelar marcante desvio da regressão linear, não por ser inerentemente irregular, mas por mostrar um padrão de resposta diferente da maioria do grupo com o qual está sendo comparado (Knight, 1970; Easton & Clements, 1973; Westcott, 1986).

Knight (1970) investiga a forma como as relações entre o rendimento e características ambientais individuais são expressas através da técnica de regressão linear simples. Ele comenta que é essencial reconhecer que a síntese das características ambientais provida pelo índice de ambiente corresponde a uma transformação que lineariza a resposta média dos genótipos. Como conseqüência, podem resultar distorções e confundimento de efeitos de características individuais do ambiente sobre o rendimento, que devem ser levadas em conta em interpretações biológicas.

Dessa forma, a utilização da técnica de regressão linear e, especialmente, a interpretação da variância do desvio da regressão como parâmetro de estabilidade fenotípica devem ser efetuadas com cautela. Em última instância, será necessário determinar as principais características ambientais limitantes que influenciam a resposta e pesquisar suas implicações sobre esta.

O fato de o uso do índice de ambiente como variável preditora induzir o ajuste perfeito de uma linha reta com declividade igual a 1 global para todos os genótipos não implica, necessariamente, que a relação deva ser linear para cada genótipo em particular. De modo geral, a relação entre a resposta média individual de um genótipo e a condição ambiental representada pelo índice de ambiente pode ser não-linear (Knight, 1970; Wright, 1971; Witcombe & Whittington, 1971; Easton & Clements, 1973; Freeman, 1973). Eberhart & Russell (1966) e Verma et al. (1978) sugerem a possibilidade de curva de resposta quadrática. Okuno & Okuno (1975) acrescentam que a aproximação dessas relações através de curvas polinomiais pode requerer termos adicionais de graus superiores a um.

Verma et al. (1978) argumentam que em quase todas as espécies cultivadas há genótipos apropriados para ambientes pobres e para ambientes ricos. Ademais, em uma coleção suficientemente grande de genótipos é possível imaginar genótipos com diferentes níveis de desempenho médio e sensibilidade ao ambiente nestes dois conjuntos contrastantes de ambientes. Os seguintes tipos de comportamento de genótipos, entre outros, devem ocorrer: taxa de resposta à variação de ambiente baixa sob condições ambientais desfavoráveis e elevada em ambientes favoráveis, igual taxa de resposta em ambientes desfavoráveis e favoráveis, e taxa de resposta elevada em ambientes desfavoráveis, e baixa, em ambientes favoráveis.

O método de regressão linear simples de Finlay & Wilkinson (1963) e Eberhart & Russell (1966) não permite a distinção desses diferentes padrões de resposta e, portanto, não é capaz de identificar genótipos com as características desejáveis, isto é, genótipos que sejam responsivos a ambientes favoráveis ou melhorados e mantenham rendimentos razoáveis em ambientes adversos.

Com esse argumento, Silva & Barreto (1985) propõem representar a resposta de um genótipo à gama de ambientes por um gráfico composto de dois segmentos de reta conectados no ponto correspondente ao índice de ambiente nulo.

Este método tem sido utilizado e discutido na literatura (Riede & Barreto, 1985; Fonseca Júnior, 1987; Chaves et al., 1989; Cruz et al., 1989; Storck,

1989; Verneti et al., 1990; Souza et al., 1991; Duarte & Zimmermann, 1992; Storck & Vencovsky, 1992; Vencovsky & Barriga, 1992). Entretanto, seu potencial analítico parece não ter sido adequadamente explorado. Este artigo tem como propósito a exposição dos fundamentos deste método e sua abordagem como uma extensão do método de regressão linear simples para proporcionar maior flexibilidade para a caracterização dos distintos comportamentos das respostas de genótipos à variação de ambiente. São discutidas e propostas questões que suscitam pesquisas. A aplicação do método será ilustrada em artigo que segue.

MÉTODO DE REGRESSÃO LINEAR SEGMENTADA

Modelo estatístico

Postula-se que o rendimento médio observado do i -ésimo genótipo no j -ésimo ambiente, denotado por y_{ij} , é expresso pela seguinte equação:

$$y_{ij} = m + c_i + a_j + ac_{ij} + \varepsilon_{ij}, \quad i=1,2,\dots,C, j=1,2,\dots,A,$$

onde: m é a média geral dos C genótipos na população representada pelos A ambientes, c_i é o efeito diferencial, fixo, do i -ésimo genótipo, a_j é o efeito aleatório do j -ésimo ambiente, ac_{ij} é o efeito aleatório da interação referente à combinação do i -ésimo genótipo com o j -ésimo ambiente, e ε_{ij} é o desvio entre os rendimentos médios observado e esperado do genótipo i no ambiente j , suposto aleatório e satisfazer as propriedades de distribuição usuais do componente do erro, ou seja, homogeneidade de variância, ausência de correlação e normalidade.

Para maior simplicidade da exposição que segue, supõe-se que a estrutura dos dados é balanceada, ou seja, que os C genótipos são comuns para os A ambientes e que o número de repetições dos genótipos é comum para todos os genótipos em todos os ambientes. O método pode ser estendido para situação mais genérica.

Silva & Barreto (1985) propõem separar do efeito da interação genótipo \times ambiente um componente atribuível à sua relação sistemática com o ambiente, exprimindo-o através da seguinte equação:

$$ac_{ij} = k_{1i}' a_j + k_{2i} \delta_j a_j + d_{ij},$$

onde: k_{1i}' e k_{2i} são constantes desconhecidas,

$$\delta_j = \begin{cases} 0, & \text{se } a_j \leq 0 \\ 1, & \text{se } a_j > 0 \end{cases}$$

e d_{ij} é o desvio entre a interação e seu componente sistemático.

Substituindo essa expressão da interação na equação do rendimento médio observado, obtém-se:

$$y_{ij} = m + c_i + a_j + k_{1i}' a_j + k_{2i} \delta_j a_j + e_{ij}, \\ = m_i + k_{1i} a_j + k_{2i} \delta_j a_j + e_{ij}, \quad i=1,2,\dots,C, j=1,2,\dots,A,$$

onde: $m_i = m + c_i$, $k_{1i} = 1 + k_{1i}'$ e $e_{ij} = \varepsilon_{ij} + d_{ij}$.

Como o efeito de ambiente a_j não é observável, ele é substituído pelo *índice de ambiente*, que, para cada ambiente, é o desvio de sua média em relação à média geral: $I_j = \bar{y}_{.j} - \bar{y}$. Dessa forma, a equação do modelo estatístico postulado para o rendimento médio do i -ésimo genótipo no j -ésimo ambiente tem a seguinte expressão:

$$y_{ij} = b_{0i} + b_{1i} I_j + b_{2i} \delta_j I_j + e_{ij}, \quad i=1,2,\dots,C, j=1,2,\dots,A,$$

onde: b_{0i} , b_{1i} e b_{2i} , $i=1,2,\dots,C$, são constantes desconhecidas, I_j , $j=1,2,\dots,A$, são valores conhecidos e e_{ij} , $i=1,2,\dots,C, j=1,2,\dots,A$, são variáveis aleatórias.

A correspondente equação para valores esperados, para cada genótipo, pode ser alternativamente escrita como a composição de duas equações:

$$E(Y_i) = \begin{cases} b_{0i} + b_{1i} I, & I \leq 0 \\ b_{0i} + (b_{1i} + b_{2i}) I, & I > 0 \end{cases}$$

$i=1,2,\dots,C$.

Essa é a expressão analítica do gráfico de dois segmentos de reta conectados no ponto $I=0$, onde: b_{1i} é a declividade do segmento de reta no subespaço correspondente a índices de ambiente negativos (*ambientes desfavoráveis*); $b_{1i}+b_{2i}$ é a declividade do segmento de reta no subespaço de índices de ambiente positivos (*ambientes favoráveis*); e b_{2i} é a variação da declividade entre esses dois subespaços. A representação gráfica dessa reta segmentada é apresentada na Fig. 1.

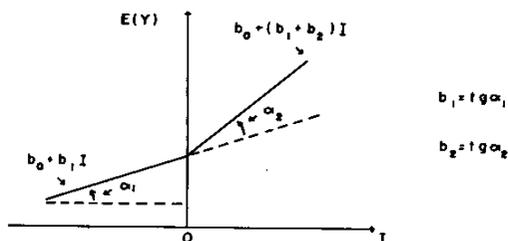


FIG. 1. Gráfico da equação do modelo linear segmentado de Corrêa da Silva & Barreto (1985).

Observe-se que esse modelo linear segmentado reduz-se ao modelo de regressão linear simples (Finlay & Wilkinson, 1963; Eberhart & Russell, 1966) quando $b_{2i} = 0$. Ademais, as retas segmentadas para os C genótipos apresentam propriedade semelhante à dessas retas de regressão, ou seja, a declividade do gráfico médio para os C genótipos é igual a um. Isso significa que:

$$\bar{b}_{1.} = \frac{1}{C} \sum_{i=1}^C b_{1i} = 1 \quad \text{e} \quad \bar{b}_{2.} = \frac{1}{C} \sum_{i=1}^C b_{2i} = 0.$$

O modelo linear segmentado é um modelo de regressão linear múltipla conjunta:

$$y_{ij} = b_{0i} + b_{1i} x_{1j} + b_{2i} x_{2j} + e_{ij}, \quad i=1,2,\dots,C, \\ j=1,2,\dots,A,$$

onde: b_{0i} , b_{1i} e b_{2i} são a interseção e os coeficientes de regressão parciais para o i -ésimo genótipo, $x_{1j}=I_j$ e $x_{2j}=\delta_j I_j$ ($\delta_j=0$ se $I_j \leq 0$, $=1$ se $I_j > 0$).

Essa é uma forma particular do modelo linear geral, cuja equação é expressa, em notação matricial, por:

$$y = Xb + e.$$

Pelo modelo de regressão linear segmentada (Silva & Barreto, 1985), a adaptabilidade de um genótipo é caracterizada pelos seguintes parâmetros: Pelo modelo de regressão linear segmentada (Silva & Barreto, 1986), a adaptabilidade de um genótipo é caracterizada pelos seguintes parâmetros: rendimento médio ($b_{0i} + b_{2i} \bar{x}_2$, onde \bar{x}_2 é a média dos índices de ambiente positivos), taxa de resposta nos ambientes desfavoráveis (b_{1i}) e taxa de resposta nos ambientes favoráveis ($b_{1i} + b_{2i}$). O genótipo desejável é aquele a que correspondem elevado rendimento médio, taxa de resposta baixa nos ambientes desfavoráveis e elevada, nos ambientes favoráveis.

As combinações de taxa de resposta acima da média, taxa de resposta média e taxa de resposta abaixo da média nos ambientes desfavoráveis e nos ambientes favoráveis definem as nove categorias de sensibilidade de resposta apresentadas na Tabela 1. Em geral, as categorias 1 e 2 são as mais desejáveis, e as categorias 8 e 9, as mais indesejáveis. As categorias 1, 2 e 3, nesta ordem, são as preferíveis para ambientes desfavoráveis; as categorias 1, 4 e 7, nesta ordem, são as preferíveis para ambientes favoráveis.

Por sua vez, as combinações dessas 9 distintas categorias de sensibilidade de resposta com as três classes de resposta média, ou seja, resposta acima da média, resposta média e resposta abaixo da média, definem 27 categorias de adaptabilidade. A distinção e ordenação de tal número elevado de categorias pode se tornar extremamente complexa. Comumente, entretanto, os genótipos se classificam em apenas um subconjunto dessas categorias, de modo que a classificação das distintas categorias de adaptabilidade e sua interpretação não será, em geral, tão complexa.

As variâncias dos desvios da regressão ($\sigma_{d_i}^2$, $i=1,2,\dots,C$) devem ser consideradas com cautela na avaliação da estabilidade fenotípica. Desvios de regressão significativos podem ser uma indicação de que o modelo linear postulado não é completamente apropriado para a caracterização da estabilidade fenotípica e adaptabilidade. As origens de desvios elevados devem ser investigadas através da inspeção das características ambientais e das de-

mais condições e ocorrências durante a condução do experimento. Uma análise de resíduos (Draper & Smith, 1981; Cook & Weisberg, 1982) pode ser útil para esses propósitos.

Estimação dos parâmetros

Os estimadores dos parâmetros b_{0i} , b_{1i} e b_{2i} do modelo de regressão linear segmentada, para cada genótipo ($i=1,2,\dots,C$), derivados pelo método dos quadrados mínimos (Snedecor & Cochran, 1980; Draper & Smith, 1981, por exemplo), são a solução do sistema de equações normais:

$$\begin{bmatrix} A & 0 & \sum_{j=1}^A x_{2j} \\ 0 & \sum_{j=1}^A x_{1j}^2 & \sum_{j=1}^A x_{2j}^2 \\ \sum_{j=1}^A x_{2j} & \sum_{j=1}^A x_{2j}^2 & \sum_{j=1}^A x_{2j}^2 \end{bmatrix} \begin{bmatrix} \hat{b}_{0i} \\ \hat{b}_{1i} \\ \hat{b}_{2i} \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^A y_{ij} \\ \sum_{j=1}^A x_{1j} y_{ij} \\ \sum_{j=1}^A x_{2j} y_{ij} \end{bmatrix}$$

ou seja:

$$\hat{b}_{1i} = \frac{1}{D} [\text{SQ}x_2 \text{SP}x_1 y(i) - \text{SP}x_1 x_2 \text{SP}x_2 y(i)],$$

$$\hat{b}_{2i} = \frac{1}{D} [\text{SQ}x_1 \text{SP}x_2 y(i) - \text{SP}x_1 x_2 \text{SP}x_1 y(i)] \text{ e}$$

$$\hat{b}_{0i} = \bar{y}_i - \hat{b}_{2i} \bar{x}_2,$$

onde:

$$D = \text{SQ}x_1 \text{SQ}x_2 - [\text{SP}x_1 x_2]^2,$$

$$\text{SQ}x_1 = \sum_{j=1}^A x_{1j}^2$$

$$\text{SQ}x_2 = \sum_{j=1}^A x_{2j}^2 - \frac{1}{A} \left(\sum_{j=1}^A x_{2j} \right)^2,$$

$$\text{SP}x_1 x_2 = \sum_{j=1}^A x_{2j}^2,$$

$$\text{SP}x_1 y(i) = \sum_{j=1}^A x_{1j} y_{ij} \text{ e}$$

$$\text{SP}x_2 y(i) = \sum_{j=1}^A x_{2j} y_{ij} - \frac{1}{A} \left(\sum_{j=1}^A x_{2j} \right) \left(\sum_{j=1}^A y_{ij} \right)$$

Os estimadores das variâncias e da covariância dos estimadores dos parâmetros b_{1i} e b_{2i} , para cada genótipo ($i=1,2,\dots,C$), são:

$$\hat{\text{Var}}(\hat{b}_{1i}) = \frac{\text{SQ}x_2}{D} s_{Y.X_1 X_2}^2(i)$$

$$\hat{\text{Var}}(\hat{b}_{2i}) = \frac{\text{SQ}x_1}{D} s_{Y.X_1 X_2}^2(i) \text{ e}$$

$$\hat{\text{Cov}}(\hat{b}_{1i}, \hat{b}_{2i}) = -\frac{\text{SP}x_1 x_2}{D} s_{Y.X_1 X_2}^2(i),$$

onde $s_{Y.X_1 X_2}^2(i)$ é o resíduo da regressão para o i -ésimo genótipo, isto é, o QM Resíduo(i).

O estimador da variância do desvio da regressão para o i -ésimo genótipo é:

$$s_{d_i}^2 = \frac{1}{A-3} \sum_{j=1}^A \hat{e}_{ij}^2 - \frac{s_c^2}{R} = \text{QM Resíduo}(i) - \text{QM}$$

Erro Combinado,

onde o QM Erro Combinado é a estimativa da variância de uma média de genótipo em um ambiente (s_c^2 é a média dos quadrados médios dos erros dos A ambientes e R é o número de repetições dos genótipos em cada um dos ambientes).

As estimativas dos parâmetros de adaptabilidade b_{0i} , b_{1i} e b_{2i} e das correspondentes variâncias e covariâncias e a estimativa da variância do desvio

da regressão, para cada genótipo, podem ser obtidas através de procedimentos de análise de regressão linear disponíveis em muitos programas computacionais para análise estatística.

Testes de hipóteses

O modelo linear segmentado propicia testes de diversas hipóteses sobre os parâmetros b_{1i} e b_{2i} , $i=1,2,\dots,C$. Essas hipóteses, listadas a seguir, referem-se à adequabilidade do modelo linear segmentado para a análise de adaptabilidade e à caracterização comparativa da adaptabilidade dos C genótipos.

1. Igualdade das taxas de resposta dos C genótipos nos ambientes desfavoráveis e nos ambientes favoráveis (ou seja, igualdade das declividades dos segmentos de reta dos C genótipos em cada um desses dois subespaços de ambientes):

$$\begin{cases} H_{01}: b_{1i} = b_1, i = 1, 2, \dots, C \text{ (constante)} \text{ e } b_{2i} = b_2, i = 1, 2, \dots, C \text{ (constante)} \\ H_{a1}: b_{1i} \neq b_{1j} \text{ e/ou } b_{2i} \neq b_{2j}, \text{ para pelo menos um par de genótipos } i \text{ e } j \end{cases}$$

Esta é a hipótese de igualdade simultânea dos coeficientes de regressão parciais b_{1i} , $i=1,2,\dots,C$, e b_{2i} , $i=1,2,\dots,C$.

A rejeição da hipótese H_{01} é condição *sine qua non* para a aplicabilidade do método de regressão linear segmentada. A aplicação desse método quando a hipótese H_{01} é aceita implicaria o ajustamento de retas com declividade comum $b=1$ para todos os C genótipos. Nesta circunstância, esse método não seria capaz de detectar distinções entre os genótipos quanto à adaptabilidade. Neste caso, as hipóteses listadas a seguir perdem interesse.

2. Igualdade das diferenças de taxa de resposta entre ambientes desfavoráveis e favoráveis para os C genótipos:

$$\begin{cases} H_{02}: b_{2i} = b_2, i = 1, 2, \dots, C \text{ (constante)} \\ H_{a2}: b_{2i} \neq b_{2j}, \text{ para pelo menos um par de genótipos } i \text{ e } j \end{cases}$$

Esta é a hipótese de igualdade dos coeficientes de regressão parciais b_{2i} , $i=1,2,\dots,C$.

O teste desta hipótese é crucial para a decisão referente à forma do modelo linear a adotar. Como $b_{2i} = b_2$, $i=1,2,\dots,C$, implica $b_{2i} = 0$, $i=1,2,\dots,C$, a aceitação da hipótese H_{02} é indicação de que a relação entre rendimento e índice de ambiente pode ser apro-

priamente representada por um modelo de regressão linear simples. A rejeição desta hipótese de nulidade indica que o modelo de regressão linear segmentada completo é o mais adequado.

Na primeira situação, procede-se a aplicação do método de regressão linear simples de Finlay & Wilkinson (1963) e Eberhart & Russell (1966). Neste caso, as hipóteses listadas a seguir não têm sentido.

3. Igualdade das taxas de resposta dos C genótipos nos ambientes desfavoráveis:

$$\begin{cases} H_{03}: b_{1i} = b_1, i = 1, 2, \dots, C \text{ (constante)} \\ H_{a3}: b_{1i} \neq b_{1j}, \text{ para pelo menos um par de genótipos } i \text{ e } j \end{cases}$$

4. Igualdade das taxas de resposta dos C genótipos nos ambientes favoráveis:

$$\begin{cases} H_{04}: b_{1i} + b_{2i} = b, i = 1, 2, \dots, C \text{ (constante)} \\ H_{a4}: b_{1i} + b_{2i} \neq b_{1j} + b_{2j}, \text{ para pelo menos um par de genótipos } i \text{ e } j \end{cases}$$

5. Igualdade das taxas de resposta entre ambientes desfavoráveis e favoráveis, para cada genótipo:

$$(i) \begin{cases} H_{05}: b_{2i} = 0 \\ H_{a5}: b_{2i} \neq 0 \end{cases}$$

$i=1,2,\dots,C$.

6. Nulidade da taxa de resposta nos ambientes desfavoráveis, para cada genótipo:

$$(i) \begin{cases} H_{06}: b_{1i} = 0 \\ H_{a6}: b_{1i} \neq 0 \end{cases}$$

$i=1,2,\dots,C$.

7. Nulidade da taxa de resposta nos ambientes favoráveis, para cada genótipo:

$$(i) \begin{cases} H_{07}: b_{1i} + b_{2i} = 0 \\ H_{a7}: b_{1i} + b_{2i} \neq 0 \end{cases}$$

$i=1,2,\dots,C$.

8. Taxa de resposta nos ambientes desfavoráveis igual à taxa de resposta média dos genótipos, para cada genótipo particular:

$$(i) \begin{cases} H_{08}: b_{1i} = 1 \\ H_{a8}: b_{1i} \neq 1 \end{cases}$$

$$i=1,2,\dots,C.$$

9. Taxa de resposta nos ambientes favoráveis igual à taxa de resposta média dos genótipos, para cada genótipo particular:

$$(i) \begin{cases} H_{09}: b_{1j} + b_{2i} = 1 \\ H_{a9}: b_{1j} + b_{2i} \neq 1 \end{cases}$$

$$i=1,2,\dots,C.$$

Os testes dos dois últimos conjuntos de hipóteses propiciam a classificação dos genótipos nas nove categorias de sensibilidade de resposta caracterizadas na Tabela 1.

TABELA 1. Distintas categorias de sensibilidade de resposta de genótipos caracterizadas pelo modelo de regressão linear segmentada.

Categoria	Ambientes desfavoráveis	Ambientes favoráveis
1	$b_1 < 1$	$b_1 + b_2 > 1$
2	$b_1 < 1$	$b_1 + b_2 = 1$
3	$b_1 < 1$	$b_1 + b_2 < 1$
4	$b_1 = 1$	$b_1 + b_2 > 1$
5	$b_1 = 1$	$b_1 + b_2 = 1$
6	$b_1 = 1$	$b_1 + b_2 < 1$
7	$b_1 > 1$	$b_1 + b_2 > 1$
8	$b_1 > 1$	$b_1 + b_2 = 1$
9	$b_1 > 1$	$b_1 + b_2 < 1$

Todas essas hipóteses são casos particulares da forma geral das hipóteses lineares referentes aos parâmetros do modelo linear geral $y = Xb + e$, ou seja:

$$\begin{cases} H_0: c'b = 0 \\ H_a: c'b \neq 0 \end{cases}$$

Essas hipóteses podem ser testadas pelo procedimento geral para teste de hipótese do modelo linear geral, através da seguinte estatística:

$$F = \frac{[SQ \text{ Resíduo}(H_0) - SQ \text{ Resíduo}]/v_1}{SQ \text{ Resíduo}/v_2},$$

onde o numerador é a soma de quadrados atribuível à hipótese $H_0: c'b = 0$, ou seja, SQ Hipótese H_0 ; SQ Resíduo é a soma de quadrados do resíduo correspondente ao ajustamento do modelo $y = Xb + e$, com v_2 graus de liberdade; SQ Resíduo(H_0) é a soma de quadrados do resíduo para o modelo sob as condições impostas pelo conjunto de equações $c'b = 0$; e $v_1 = GL$ Hipótese H_0 é o número de condições independentes sobre o conjunto de parâmetros b impostas por estas equações, estabelecidas pela hipótese H_0 . Essa estatística tem distribuição F com v_1 e v_2 graus de liberdade (Draper & Smith, 1981).

Esses testes de hipóteses podem ser efetuados, facilmente, com o uso de "pacotes" de análise estatística que disponham de procedimento para a análise do modelo linear geral, como o procedimento GLM do SAS (SAS Institute Inc., 1985).

A SQ Resíduo do denominador comum das estatísticas F para os testes das hipóteses 1, 2, 3 e 4 é a soma de quadrados do resíduo do ajustamento do modelo de regressão linear segmentada conjuntamente para os C genótipos, ou seja:

$$E(Y_{ij}) = b_{0i} + b_{1i} x_{1j} + b_{2i} x_{2j}, \quad i=1,2,\dots,C, j=1,2,\dots,A.$$

As SQ Resíduo(H_0) para os testes dessas quatro hipóteses são providas pelos ajustamentos desse modelo conjunto sob as condições impostas aos parâmetros pelas correspondentes hipóteses H_0 , ou seja, pelos ajustamentos dos seguintes modelos:

$$\text{Hipótese 1: } E(Y_{ij}) = b_{0i} + b_1 x_{1j} + b_2 x_{2j}, \quad i=1,2,\dots,C, j=1,2,\dots,A;$$

$$\text{Hipótese 2: } E(Y_{ij}) = b_{0i} + b_{1i} x_{1j} + b_2 x_{2j}, \quad i=1,2,\dots,C, j=1,2,\dots,A;$$

$$\text{Hipótese 3: } E(Y_{ij}) = b_{0i} + b_1 x_{1j} + b_{2i} x_{2j}, \quad i=1,2,\dots,C, j=1,2,\dots,A;$$

$$\text{Hipótese 4: } E(Y_{ij}) = b_{0i} + b_{1i} (x_{1j} - x_{2j}) + b_2 x_{2j}, \quad i=1,2,\dots,C, j=1,2,\dots,A.$$

A SQ Resíduo comum das estatísticas F para os testes das hipóteses 5, 6, 7, 8 e 9, para cada um dos genótipos, é provida pelo ajustamento do modelo

de regressão linear segmentada separadamente para cada um dos genótipos. As SQ Resíduo(H_0) para os testes dessas hipóteses são providas pelos ajustamentos, separados para os C genótipos, desse modelo sob as condições impostas aos parâmetros b_{1j} e b_{2j} pelas correspondentes hipóteses H_0 , ou seja, pelos ajustamentos dos seguintes modelos:

Hipóteses 5: $E(Y_{ij}) = b_{0i} + b_{1j} x_{1j}$, $j=1,2,\dots,A$ ($i=1,2,\dots,C$);

Hipóteses 6: $E(Y_{ij}) = b_{0i} + b_{2j} x_{2j}$, $j=1,2,\dots,A$ ($i=1,2,\dots,C$);

Hipóteses 7: $E(Y_{ij}) = b_{0i} + b_{1j} (x_{1j} - x_{2j})$, $j=1,2,\dots,A$ ($i=1,2,\dots,C$);

Hipóteses 8: $E(Y_{ij}) - x_{1j} = b_{0i} + b_{2j} x_{2j}$, $j=1,2,\dots,A$ ($i=1,2,\dots,C$);

Hipóteses 9: $E(Y_{ij}) - x_{2j} = b_{0i} + b_{1j} (x_{1j} - x_{2j})$, $j=1,2,\dots,A$ ($i=1,2,\dots,C$).

Um procedimento conveniente para os testes das hipóteses 1 e 2 é a decomposição da soma de quadrados da interação genótipo x ambiente na análise da variação conjunta de ambientes, conforme indicado a seguir.

A hipótese 1 pode ser testada semelhantemente ao procedimento para o método de regressão linear simples. A SQ Genótipo x Ambiente, com $(C-1)(A-1)$ graus de liberdade, é particionada em dois componentes ortogonais - um atribuível à heterogeneidade das linhas de regressão segmentadas (ou seja, desigualdade dos coeficientes de regressão parciais b_{1i} , $i=1,2,\dots,C$, e b_{2i} , $i=1,2,\dots,C$), denotado por SQ Heter. $R(X_1, X_2)$, com $2(C-1)$ graus de liberdade, e o outro exprimindo os resíduos, SQ Resíduo ou SQ Resíduo Combinado, com $(C-1)(A-3)$ graus de liberdade. A estatística F para o teste desta hipótese é a razão dos quadrados médios desses dois componentes.

Pode-se verificar que a SQ Resíduo Combinado, com $(C-1)(A-3)$ graus de liberdade, é igual à soma das somas de quadrados dos desvios do modelo linear segmentado ajustado individualmente para cada genótipo, com $C(A-3)$ graus de liberdade. A diferença entre esses números de graus de liberdade, ou seja, $C(A-3) - (C-1)(A-3) = A-3$, são os graus de liberdade subtraídos da SQ Resíduo Combinado para levar em conta a dependência linear dos resíduos

implicada pelo ajuste perfeito da equação do modelo para as médias globais dos A ambientes. Esses A-3 graus de liberdade, aos quais corresponde soma de quadrados zero, devem ser subtraídos dos números de graus de liberdade das SQ Resíduo e das SQ Resíduo(H_0) correspondentes a cada um dos ajustamentos do modelo de regressão linear segmentada conjunto para os C genótipos.

Uma complementação da informação fornecida pelo teste da hipótese 1 referente à adequabilidade do método de regressão linear segmentada é provida pela comparação dos quadrados médios do resíduo combinado e do erro combinado, através da seguinte estatística: $F = \text{QM Resíduo Combinado} / \text{QM Erro Combinado}$.

As implicações dos resultados desses dois testes de hipóteses referentes à adequabilidade do método de regressão linear segmentada são semelhantes às descritas para o método de regressão linear simples.

O teste da hipótese 2 pode ser procedido através da decomposição da SQ Heter. $R(X_1, X_2)$ em dois componentes ortogonais - um atribuível à heterogeneidade das retas (não segmentadas) ajustadas para todos os ambientes (ou seja, desigualdade dos coeficientes de regressão b_{1i} , $i=1,2,\dots,C$, ignorando X_2), denotado por SQ Heter. $R(X_1)$, com C-1 graus de liberdade, e o outro atribuível à variação adicional devida à heterogeneidade das alterações de declividade das retas segmentadas entre os ambientes desfavoráveis e favoráveis (ou seja, desigualdade dos coeficientes de regressão parciais b_{2i} , $i=1,2,\dots,C$, após ajustamento para X_1), denotado por SQ Heter. $R(X_2|X_1)$, também com C-1 graus de liberdade.

Os testes das hipóteses 1 e 2 podem ser efetuados com o auxílio de uma tabela com a forma da Tabela 2, que especifica os graus de liberdade e as origens

TABELA 2. Decomposição da interação genótipo x ambiente para os teste das hipóteses 1 e 2.

Fonte de variação	GL	SQ
Genotipo x Ambiente	(C-1)(A-1)	
Heter. $R(X_1, X_2)$	2(C-1)	SQ Resíduo(H_{01}) - SQ Resíduo
Heter. $R(X_1)$	C-1	SQ Resíduo(H_{01}) - SQ Resíduo(H_{02})
Heter. $R(X_2 X_1)$	C-1	SQ Resíduo(H_{02}) - SQ Resíduo
Resíduo Combinado	(C-1)(A-3)	SQ Resíduo

das somas de quadrados correspondentes às fontes de variação resultantes da decomposição sucessiva da SQ Genótipo x Ambiente.

Os testes das hipóteses 3 e 4 e dos conjuntos de hipóteses 7, 8 e 9 podem ser efetuados com o auxílio de tabelas, tal como a Tabela 3, que especifica os correspondentes graus de liberdade apropriados.

As hipóteses 5 e 6, para cada genótipo, são as hipóteses parciais referentes aos parâmetros b_{2i} e b_{1i} , respectivamente.

Comparações dos parâmetros de adaptabilidade

Se alguma das hipóteses 2, 3 e 4 revela-se significativa, podem ser de interesse comparações das taxas de resposta dos genótipos nos ambientes desfavoráveis e nos ambientes favoráveis e das alterações de taxa de resposta dos genótipos entre esses dois subespaços de ambientes. Ou seja, respectivamente, comparações dos parâmetros de adaptabilidade b_{1i} , $i=1,2,\dots,C$, $b_{1i}+b_{2i}$, $i=1,2,\dots,C$, e b_{2i} , $i=1,2,\dots,C$. Tais comparações podem ser efetuadas pelo teste da diferença mínima significativa (teste *dms*) de Fisher, ou outro teste de comparações múltiplas apropriado.

As diferenças mínimas significativas para as comparações múltiplas dos parâmetros b_{1i} , $i=1,2,\dots,C$, b_{2i} , $i=1,2,\dots,C$, e $b_{1i}+b_{2i}$, $i=1,2,\dots,C$, pelo teste *dms* de Fisher, no nível de significância α , são dadas pelas seguintes expressões:

$$b_{1i} - b_{1i'}: \text{dms} = t(v, \alpha) \sqrt{2 \hat{\text{Var}}(\hat{b}_{1i})}$$

$$b_{2i} - b_{2i'}: \text{dms} = t(v, \alpha) \sqrt{2 \hat{\text{Var}}(\hat{b}_{2i})}$$

$$b_{1i} + b_{2i} - b_{1i'} - b_{2i'}: \text{dms} = t(v, \alpha) \sqrt{2 \hat{\text{Var}}(\hat{b}_{1i} + \hat{b}_{2i})}$$

TABELA 3. Tabela para os testes das hipóteses 3 e 4 e dos conjuntos de hipóteses 7, 8 e 9.

Fonte de variação	GL	
	Hip. 3 e 4	Hip. 7, 8 e 9
Hipótese H_0	C-1	1
Resíduo	(C-1)(A-3)	A-3
Resíduo(H_0)	(C-1)(A-2)	A-2

onde $t(v; \alpha)$ é o ponto $\alpha\%$ superior da distribuição t de Student para $v = (C-1)(A-3)$ graus de liberdade.

Se as estimativas das variâncias e covariâncias de \hat{b}_{1i} e \hat{b}_{2i} para os C genótipos revelarem igualdades das correspondentes variâncias e covariâncias populacionais, será mais apropriado a utilização das correspondentes estimativas médias comuns para os C genótipos.

DISCUSSÃO

O método de regressão linear simples de Finlay & Wilkinson (1963) e Eberhart & Russell (1966) condiciona uma relação linear exata entre a resposta média de ambiente e o índice de ambiente. Se os genótipos em consideração apresentam semelhantes padrões de resposta às características ambientais, aquela condição implica a linearidade das relações entre as respostas individuais dos genótipos e o índice de ambiente. Por essa razão, a técnica de regressão linear simples tem sido aplicada com sucesso em muitos estudos de estabilidade fenotípica e adaptabilidade de plantas cultivadas.

Entretanto, em geral, relações individuais não-lineares podem ocorrer, em consequência de diferenças entre os padrões de resposta dos genótipos. Tais diferenças de comportamento das respostas de genótipos não são detectadas pelo método de regressão linear simples.

Em algumas situações, uma aproximação razoável de uma relação não-linear pode ser provida por uma função polinomial de grau baixo, ou por uma função linear segmentada, com pequeno número de segmentos.

O modelo linear segmentado proposto por Silva & Barreto (1985) aproxima a relação entre a resposta média de cada genótipo e o índice de ambiente por um gráfico composto de dois segmentos de reta conectados no ponto correspondente ao índice de ambiente nulo. Ele se reduz ao modelo de regressão linear simples quando a diferença das declividades desses dois segmentos de reta é nula para cada genótipo. Nessas circunstâncias, o modelo de regressão linear segmentada é uma extensão do modelo de regressão linear simples que leva em conta a possibilidade de diferenças de sensibilidade de respos-

ta dos genótipos entre ambientes desfavoráveis e favoráveis.

O método de regressão linear segmentada tem a mesma conveniência do método de regressão linear simples de propiciar a caracterização da adaptabilidade e estabilidade fenotípica através de poucos parâmetros de um modelo linear de fácil aplicação. Entretanto, também é passível de críticas e restrições semelhantes às que têm sido apresentadas ao método de regressão linear simples, e das críticas e restrições gerais referentes a essas pesquisas. Essas críticas e restrições são discutidas a seguir.

1) Sintetização das características ambientais através do índice de ambiente.

A relativa simplicidade do método de regressão linear decorre, justamente, da síntese das características ambientais propiciada pelo índice de ambiente. Entretanto, as seguintes consequências devem ser avaliadas e levadas adequadamente em conta pelo pesquisador:

a) A caracterização da adaptabilidade e estabilidade fenotípica propiciada é relativa ao conjunto dos genótipos incluídos no experimento. Como a produtividade do ambiente é expressa pelo próprio desempenho médio dos genótipos no experimento, não há qualquer segurança de que um genótipo manterá o mesmo comportamento quando avaliado com um diferente conjunto de genótipos (Easton & Clements, 1973; Lin et al., 1986).

b) A representação das influências de características ambientais sobre o rendimento pelo índice de ambiente pode conduzir a interpretações enganosas. Knight (1970) salienta que as seguintes consequências devem ser levadas em conta em interpretações biológicas: i) Respostas iguais correspondentes a valores de uma mesma característica ambiental situados abaixo e acima do ótimo são justapostas; ii) diferentes características limitantes (geada e seca, por exemplo) podem resultar em iguais respostas médias baixas; entretanto, é pouco provável que genótipos tenham semelhantes ordenações com relação a essas distintas características; iii) comparações de genótipos com diferenças essenciais de amplitudes totais do ciclo vegetativo ou de fases de crescimento podem ser enganosas; iv) desvios da regressão de genótipos de um subconjunto de genótipos que determina as médias ambientais são usualmen-

te consideravelmente menores do que os desvios da regressão dos demais genótipos.

Para superar essas restrições devem ser desenvolvidos e postos à disposição dos pesquisadores modelos e métodos que relacionem o rendimento mais diretamente com as características do ambiente. Instrumentos que facilitem o registro de informações sobre essas características também devem se tornar disponíveis. Enquanto essa abordagem não for viável, a expressão da produtividade ambiental pelo rendimento médio dos genótipos deve continuar a ser explorada, juntamente com a utilização de dados de características ambientais relevantes.

2) Suposição de função de resposta linear para a relação entre rendimento e índice de ambiente.

Críticas nesse contexto referem-se a violações de pressuposições da análise de regressão linear e à inadequabilidade da forma de resposta linear para exprimir a relação entre rendimento e índice de ambiente (Baker, 1969; Freeman & Perkins, 1971; Hardwick & Wood, 1972; Freeman, 1973; Easton & Clements, 1973; Hill, 1975; Westcott, 1986). Elas são as seguintes:

a) As variáveis preditoras na análise de regressão linear são funções do índice de ambiente; consequentemente, elas são sujeitas a erro.

b) As médias de genótipos contribuem para o índice de ambiente e, portanto, não são estatisticamente independentes deste índice.

c) As estimativas dos parâmetros que caracterizam a estabilidade e adaptabilidade não são independentes. Desta forma, as estimativas desses parâmetros contêm informações redundantes.

d) Ajustamentos de regressão podem ser indevidamente influenciados pelo desempenho em relativamente poucos ambientes.

e) A proporção da soma de quadrados da interação genótipo x ambiente levada em conta pelo modelo estatístico linear pode ser pequena, o que significa que a relação linear pode ser inadequada para exprimir a variação da resposta dos genótipos à produtividade do ambiente.

f) Na prática, parâmetros de regressão podem falhar na caracterização e distinção de padrões de adaptabilidade e estabilidade de genótipos.

A consequência da violação que origina a primeira crítica é a tendenciosidade das estimativas dos coeficientes de regressão. Entretanto, não resultará qualquer tendenciosidade séria, se o experimento incluir um número razoável de genótipos e a amplitude de ambientes for tal que o quadrado médio de ambiente seja significativamente maior que o quadrado médio do erro (Hardwick & Wood, 1972; Hill, 1975).

Com referência à segunda crítica, Freeman (1973) reconhece que esse procedimento é perfeitamente válido, desde que os efeitos de genótipo e de ambiente sejam considerados fixos. Isso significa que inferências válidas se restringem aos conjuntos de genótipos e de ambientes presentes no experimento. A extensão de inferências referentes ao comportamento de genótipos a ambientes não incluídos no experimento será válida na medida em que esses ambientes forem semelhantes aos presentes no experimento (Hill, 1975).

Essas duas primeiras críticas poderão ser superadas pela adoção de modelo estatístico com suposição alternativa referente ao termo que representa o efeito de ambiente, tal como: i) efeito fixo, constituindo um parâmetro adicional do modelo, ou ii) variável sujeita a erro.

A crítica c refere-se a uma propriedade usual de inferências referentes a parâmetros de modelos de regressão linear múltipla.

A restrição d ressalta a importância da escolha dos ambientes e da investigação de características ambientais de influências relevantes sobre a resposta. Uma análise de resíduos aplicada aos desvios da regressão pode ser útil para detectar ambientes que demandem exame mais cuidadoso, propiciar indicação de possíveis fontes de variação sistemáticas não levadas em conta pelo modelo estatístico, e sugerir indicações de alterações dos ambientes de condução do experimento na continuação do programa de melhoramento genético.

As críticas e e f salientam, apenas, que o método de regressão linear segmentada, assim como o método de regressão linear simples, não pode ser encarado pelo melhorista como uma panacéia. O método de regressão linear simples tem-se revelado útil, em muitas situações, para descrever o comportamento de genótipos em uma gama de ambientes. É de

esperar que o método de regressão linear segmentada, que o estende, possa ampliar a utilidade da abordagem de regressão linear para abranger outras situações, com a mesma conveniente habilidade para reduzir complexas interações a um conjunto ordenado de padrões de respostas. Entretanto, como qualquer outra técnica de modelagem estatística, ele também pode falhar. Muito freqüentemente, a razão biológica para essa falha poderá ser elucidada pela manipulação criteriosa dos dados do experimento com o recurso de informações sobre características ambientais.

Dessa forma, a abordagem de regressão linear deve ser considerada como uma adição aos recursos de técnicas do melhorista que, se usados corretamente, devem facilitar o processo de decisão em um programa particular de melhoramento genético.

3) Escolha do índice de ambiente zero para frente entre os ambientes desfavoráveis e favoráveis.

A arbitrariedade dessa escolha é passível de discussão. Ela pode ser apropriada nas situações em que a distribuição dos índices de ambiente seja razoavelmente simétrica. No caso de distribuição assimétrica, a mediana poderá ser preferível à média para a caracterização do centro da distribuição. Uma outra opção alternativa é a adoção do ponto correspondente à conexão dos dois segmentos de reta como um parâmetro de adaptabilidade adicional. Este parâmetro poderia fornecer informação interessante referente à variação do comportamento da resposta do genótipo. Naturalmente, com essa opção, a caracterização de categorias de adaptabilidade se tornaria mais diversificada, o que implicaria maior complexidade para a análise de adaptabilidade.

4) Uso de dados de programa experimental não planejado para propósito de análise de adaptabilidade.

Em geral, o objetivo de programas de melhoramento genético vegetal é a obtenção de genótipos de elevados rendimentos médios, que sejam responsivos a características ambientais favoráveis ou melhoradas e que mantenham desempenho razoável em ambientes adversos. Portanto, a avaliação do comportamento da resposta deveria fazer parte dos critérios de avaliação do desempenho relativo dos genótipos. Entretanto, muitos desses programas não contemplam esse objetivo.

As conseqüências da ulterior execução de análise de adaptabilidade com base em dados cuja coleta não foi planejada para esse propósito são, entre outras, as seguintes:

a) Inferências estatísticas derivadas a partir de dados não gerados para tal propósito devem ser consideradas com cautela. Os estudos de adaptabilidade e estabilidade executados nessas circunstâncias devem ser mais apropriadamente encarados como descritivos e exploratórios. Seus resultados não podem ser tomados como conclusivos, mas meramente indicativos e fornecedores de subsídios para o processo de decisão do programa de melhoramento genético.

b) Os ambientes de condução do experimento podem não constituir uma amostra representativa da população de ambientes da região de interesse. Em geral, a escolha dos ambientes de condução do experimento é restrita e arbitrária, de modo que o pesquisador não logra uma representatividade dos locais e estações de cultivo da região de interesse. Conseqüentemente, as inferências do experimento para essa população terão que ser eminentemente subjetivas.

A seleção de ambientes é um dos aspectos mais críticos em muitos estudos de adaptabilidade e estabilidade fenotípicas. A distribuição dos genótipos com respeito ao rendimento médio e ao comportamento da resposta com a variação do ambiente depende da amostra particular de anos e locais em que é baseada a análise da adaptabilidade. Dessa forma, os genótipos devem ser cultivados em um número adequado de ambientes, cobrindo toda a amplitude de condições ambientais da região de interesse, para que os parâmetros de adaptabilidade forneçam informação útil (Finlay & Wilkinson, 1963; Eberhart & Russell, 1966).

A falha da representatividade desses ambientes é notória quando o índice de ambiente apresenta distribuição irregular, concentrando valores em alguns subintervalos ou revelando valores extremos isolados. Esses valores podem ter marcante influência sobre as estimativas dos parâmetros de regressão e podem conduzir a resultados enganosos (Westcott, 1986).

Um programa de melhoramento genético deve estabelecer a escolha apropriada dos locais para a

condução do experimento e seu monitoramento periódico. A análise de adaptabilidade deve constar dos critérios de avaliação de genótipos. Sua execução em cada etapa de avaliação do programa pode propiciar, também, indicações para as apropriadas alterações de locais e número de anos de execução do experimento.

AGRADECIMENTOS

Ao Dr. Amauri Almeida Machado, pela colaboração durante a fase de revisão deste artigo.

REFERÊNCIAS

- BAKER, R.J. Genotype-environment interactions in yield of wheat. *Canadian Journal of Plant Science*, v.49, p.743-751, 1969.
- CHAVES, L.J.; VENCOVSKY, R.; GERALDI, I.O. Modelo não linear aplicado ao estudo da interação de genótipos x ambiente em milho. *Pesquisa Agropecuária Brasileira*, Brasília, v.24, n.2, p.259-268, 1989.
- COOK, R.D.; WEISBERG, S. *Residuals and influence in regression*. New York: Chapman and Hall, 1982. 230p.
- CRUZ, C.D.; TORRES, R.A.A.; VENCOVSKY, R. An alternative approach to the stability analysis proposed by Silva and Barreto. *Revista Brasileira de Genética*, Ribeirão Preto, v.12, n.2, p.567-580, 1989.
- DRAPER, N. R.; SMITH, H. *Applied regression analysis*. 2. ed. New York: John Wiley & Sons, 1981. 709p.
- DUARTE, J.B.; ZIMMERMANN, M.J.O. Comparison of three methods used for the study of adaptation and phenotypic stability in the common bean (*Phaseolus vulgaris* L.). *Revista Brasileira de Genética*, Ribeirão Preto, v.15, n.1, p.125-136, 1992.
- EASTON, H.S.; CLEMENTS, R.J. The interaction of wheat genotypes with a specific factor of the environment. *Journal of Agricultural Science*, Cambridge, v.80, p.43-52, 1973.
- EBERHART, S.A.; RUSSELL, W.A. Stability parameters for comparing varieties. *Crop Science*, Madison, v.6, p.36-40, 1966.

- FINLAY, K.W.; WILKINSON, G.N. The analysis of adaptation in a plant-breeding programme. *Australian Journal of Agricultural Research*, Victoria, v.14, p.742-754, 1963.
- FONSECA JÚNIOR, N.S. Análise de estabilidade fenotípica: considerações sobre quatro metodologias. In: SIMPÓSIO DE ESTATÍSTICA APLICADA À EXPERIMENTAÇÃO AGRONÔMICA, 2., Londrina, 1987. *Anais...* Londrina: Universidade Estadual/Departamento de Matemática Aplicada, 1987. p.193-212.
- FREEMAN, G.H. Statistical methods for the analysis of genotype-environment interactions. *Heredity*, Edinburgh, v.31, n.3, p.339-354, 1973.
- FREEMAN, G.H.; PERKINS, J.M. Environmental and genotype-environmental components of variability. VIII. Relations between genotypes grown in different environments and measures of these environments. *Heredity*, Edinburgh, v.27, p.15-23, 1971.
- HARDWICK, R.C.; WOOD, J.T. Regression methods for studying genotype-environment interactions. *Heredity*, Edinburgh, v.28, p.209-222, 1972.
- HILL, J. Genotype-environment interactions - a challenge for plant breeding. *Journal of Agricultural Science*, Cambridge, v.85, p.477-493, 1975.
- KNIGHT, R. The measurement and interpretation of genotype environment interactions. *Euphytica*, v.19, p.225-235, 1970.
- LIN, C.S.; BINNS, M.R.; LEFKOVITCH, L.P. Stability analysis: where do we stand? *Crop Science*, Madison, v.26, p.894-900, 1986.
- OKUNO, T.; OKUNO, C.S. Analysis of interaction with application to the adaptability evaluation. *Bulletin of the International Statistical Institute*, v.66, book 4, p.194-201, 1975.
- PERKINS, J.M.; JINKS, J.L. Environmental and genotype-environmental components of variability. III. Multiple lines and crosses. *Heredity*, Edinburgh, v.23, p.339-356, 1968a.
- PERKINS, J.M.; JINKS, J.L. Environmental and genotype-environmental components of variability. IV. Non-linear interactions for multiple inbred lines. *Heredity*, Edinburgh, v.23, p.525-535, 1968b.
- RIEDE, C.R.; BARRETO, J.N. Stability of cultivars recommended for the north and west regions of Paraná State. *Biometrics*, v.41, n.2, p.1093, 1986.
- SAS INSTITUTE INC. *SAS User's Guide: Statistics*. Version 5. ed. Cary, NC: SAS Institute Inc., 1985. 956p.
- SHUKLA, G.K. Some statistical aspects of partitioning genotype-environmental components of variability. *Heredity*, Londres, v.29, p.237-245, 1972.
- SILVA, J.G.C. da; BARRETO, J.N. An application of segmented linear regression to the study of genotype x environment interaction. *Biometrics*, vol.41, n.4, p.1093, 1986.
- SNEDECOR, G.W.; COCHRAN, W.G. *Statistical Methods*. 7. ed. Ames: The Iowa State University Press, 1980. 507p.
- SOUZA, F.R.S.; RAMALHO, M.A.P.; OLIVEIRA, A.C.; SANS, L.M.A. Estabilidade de cultivares de milho em diferentes épocas e locais de plantio em Minas Gerais. *Pesquisa Agropecuária Brasileira*, Brasília, v.26, n.6, p.885-892, 1991.
- STORCK, L. *Modelo de regressão bi-segmentado descontínuo com erros de medida aplicado na análise de estabilidade de cultivares*. Piracicaba, ESALQ-USP, 1989. 217p. Tese de Doutorado.
- STORCK, L.; VENCOSKY, R. Novo modelo para avaliação da estabilidade de cultivares. *Pesquisa Agropecuária Brasileira*, Brasília, v.27, n.8, p.1201-1211, 1992.
- VENCOSKY, R.; BARRIGA, P. *Genética Biométrica no Fitomelhoramento*. Ribeirão Preto: Revista Brasileira de Genética, 1992. 486 p.
- VERMA, M.M.; CHAAL, G.S.; MURTY, B.R. Limitations of conventional regression analysis. A proposed modification. *Theoretical and Applied Genetics*, v.53, p.89-91, 1978.
- VERNETTI, F.J.; GASTAL, M.F.C.; ZONTA, E.P. Estabilidade fenotípica de cultivares de soja no Sudeste do Rio Grande do Sul. *Pesquisa Agropecuária Brasileira*, Brasília, v.25, n.11, p.1593-1602, 1990.
- WESTCOTT, B. Some methods of analysing genotype-environment interaction, *Heredity*, Londres, v.56, p.243-253, 1986.
- WITCOMBE, J.R.; WHITTINGTON, W.J. A study of the genotype by environmental interaction shown by germinating seeds of *Brassica napus*. *Heredity*, Londres, v.26, p.397-411, 1971.
- WRIGHT, A.J. The analysis and prediction of some two-factor interactions in grass breeding. *Journal of Agricultural Science*, Cambridge, v.76, p.301-306, 1971.
- YATES, F.; COCHRAN, W.G. The analysis of groups of experiments. *Journal of Agricultural Science*, Cambridge, v.28, p.556-580, 1938.