

Metodologia para análise de adaptabilidade e estabilidade por meio de regressão quantílica

Lais Mayara Azevedo Barroso⁽¹⁾, Moysés Nascimento⁽¹⁾, Ana Carolina Campana Nascimento⁽¹⁾, Fabyano Fonseca e Silva⁽²⁾, Cosme Damião Cruz⁽³⁾, Leonardo Lopes Bhering⁽³⁾ e Reinaldo de Paula Ferreira⁽⁴⁾

⁽¹⁾Universidade Federal de Viçosa (UFV), Departamento de Estatística, Avenida P.H. Rolfs, s/nº, CEP 36571-000 Viçosa, MG, Brasil. E-mail: lais.azevedobarroso@gmail.com, moysesnascim@ufv.br, campanaac@gmail.com ⁽²⁾UFV, Departamento de Zootecnia. E-mail: fabyanofonseca@ufv.br ⁽³⁾UFV, Departamento de Biologia Geral. E-mail: cdruz@ufv.br, leonardo.bhering@ufv.br ⁽⁴⁾Embrapa Pecuária Sudeste, Rodovia Washington Luiz, Km 234, CEP 13560-970 São Carlos, SP, Brasil. E-mail: reinaldo.ferreira@embrapa.br

Resumo – O objetivo deste trabalho foi desenvolver e validar uma metodologia de análise da adaptabilidade e da estabilidade fenotípica baseada em regressão quantílica (RQ). Para tanto, foram simulados valores fenotípicos com distribuição simétrica e com distribuição assimétrica à direita e à esquerda, com ou sem a presença de “outliers”. A metodologia proposta foi aplicada a um conjunto de dados provenientes de um experimento com 92 genótipos de alfafa (*Medicago sativa*), avaliados em 20 ambientes, e comparada às metodologias de Eberhart & Russell e de regressão não paramétrica. A metodologia da RQ proporcionou resultados iguais ou superiores aos obtidos com as metodologias alternativas avaliadas. No entanto, a ocorrência de resultados discordantes entre as metodologias evidencia a importância de se avaliar a simetria na distribuição dos valores fenotípicos. Para distribuições simétricas, na presença de “outliers”, deve-se utilizar a RQ com valor de quantil estimado (τ) em 0,50; na ausência de “outliers”, pode-se utilizar tanto a metodologia de Eberhart & Russell quanto a RQ ($\tau = 0,50$). Para distribuições assimétricas, indica-se o uso da RQ com $\tau = 0,25$, para assimetria à direita, e com $\tau = 0,75$, para assimetria à esquerda, independentemente da presença de “outliers”.

Termos para indexação: *Medicago sativa*, distribuição assimétrica, interação genótipo x ambiente, melhoramento vegetal, outliers, regressão não paramétrica.

Methodology for analysis of adaptability and stability using quantile regression

Abstract – The objective of this work was to develop and validate a methodology for analyzing phenotypic adaptability and stability based on quantile regression (QR). For this, phenotypic values were simulated with symmetrical distribution and with asymmetrical distribution to the right and to the left, with or without outliers. The proposed methodology was applied to a data set from an experiment with 92 alfalfa (*Medicago sativa*) genotypes, evaluated in 20 environments, and compared with the methodologies of Eberhart & Russell and nonparametric regression. The QR methodology provided equal or superior results, compared to the evaluated alternative methodologies. However, the occurrence of disagreeing results between methodologies evidences the importance of evaluating symmetry in the distribution of phenotypic values. For symmetric distributions with outliers, QR should be used with estimated quantile value (τ) of 0.50; in the absence of outliers, both the methodology of Eberhart & Russel and QR ($\tau = 0.50$) may be used. For asymmetric distributions, the use of RQ with $\tau = 0.25$ is suggested for asymmetry to the right, and with $\tau = 0.75$ for asymmetry to the left, regardless of the presence of outliers.

Index terms: *Medicago sativa*, asymmetrical distribution, genotype x environment interaction, plant breeding, outliers, nonparametric regression.

Introdução

Embora haja inúmeras metodologias para a análise da adaptabilidade e da estabilidade fenotípica, como as baseadas em regressão linear simples (Eberhart & Russell, 1966), múltipla (Cruz et al., 1989) e não paramétrica (Lin & Binns, 1988; Nascimento et al.,

2010), bem como em centroides (Nascimento et al., 2009a, 2009b) e em abordagem bayesiana (Nascimento et al., 2011), nenhuma delas contempla, de forma conjunta, os problemas relacionados às distribuições assimétricas dos valores fenotípicos e à presença de observações discrepantes (“outliers”).

Quanto às distribuições assimétricas, com possível presença de “outliers”, o uso das metodologias tradicionais pode resultar em recomendações incorretas. Nesse caso, as metodologias baseadas em regressão podem obter estimativas inadequadas dos valores fenotípicos, que não reflitam a verdadeira relação entre a variação ambiental e a resposta genotípica (Nascimento et al., 2010). Já as metodologias não paramétricas, como a de Lin & Binns (1988), podem obter resultados impróprios, uma vez que seus parâmetros podem ser inflacionados, o que resultaria na estimação incorreta da adaptabilidade e da estabilidade dos genótipos. Por sua vez, quando metodologias gráficas, tais como centroides múltiplos e ampliados (Nascimento et al., 2009a, 2009b), AMMI (Gauch Jr, 2006) e GGE biplot (Yan & Tinker, 2006), são utilizadas nesse contexto de distribuição assimétrica, a presença de “outliers” pode influenciar diretamente a configuração gráfica e prejudicar a análise.

Uma possível forma de tratar simultaneamente os problemas de assimetria e de presença de “outliers” é a utilização de regressão quantílica (RQ). Diferentemente dos métodos tradicionais de regressão, que utilizam a média (valor central) para explicar a relação funcional entre a variação ambiental e a resposta fenotípica, a RQ generaliza essa explicação para qualquer quantil de interesse. Embora esta metodologia venha se popularizando por meio de aplicações nas áreas de finanças (Silva & Porto Júnior, 2006), genômica (Beyerlein et al., 2011) e pecuária (Nascimento et al., 2012), entre outras, até o momento, não há relatos sobre sua utilização em estudos de adaptabilidade e estabilidade fenotípica.

O objetivo deste trabalho foi desenvolver e validar uma metodologia de análise da adaptabilidade e da estabilidade fenotípica baseada em regressão quantílica.

Material e Métodos

Foi avaliado um conjunto de dados reais de 92 genótipos de alfafa (*Medicago sativa* L.), submetidos a 20 diferentes épocas de cortes (ambientes), por meio da RQ; a metodologia utilizada foi validada pela comparação com os resultados das metodologias de Eberhart & Russell (1966) e de regressão não paramétrica.

O modelo de regressão quantílica proposto para estudos da adaptabilidade e da estabilidade

fenotípica considera “a” ambientes, “g” genótipos e “r” repetições, e pode ser definido da seguinte forma: $Y_{ij} = \beta_{0i}(\tau) + \beta_{1i}(\tau) I_j + \varepsilon_i(\tau)$, em que: $\beta_{0i}(\tau)$ é a constante da regressão; $\beta_{1i}(\tau)$ é o coeficiente da regressão; I_j é o índice ambiental codificado,

$$\left(I_j = \sum_i Y_{ij} / g - \sum_i \sum_j Y_{ij} / ga \right),$$

que representa a qualidade do ambiente; τ refere-se ao quantil estimado ($\tau \in [0,1]$); e $\varepsilon_i(\tau)$ são os erros aleatórios independentes, identicamente distribuídos.

O conceito de adaptabilidade refere-se à capacidade de os genótipos responderem de maneira vantajosa ao estímulo do ambiente. De acordo com a metodologia proposta, considerou-se que, quando: $\beta_{1i}(\tau) = 1$, os genótipos apresentam adaptabilidade ampla ou geral; $\beta_{1i}(\tau) > 1$, os genótipos apresentam adaptabilidade específica a ambientes favoráveis; e $\beta_{1i}(\tau) < 1$, os genótipos apresentam adaptabilidade específica a ambientes desfavoráveis. Já o conceito de estabilidade refere-se à capacidade de os genótipos mostrarem um comportamento previsível, em função do estímulo do ambiente. A estabilidade foi estimada por meio do coeficiente de determinação da RQ, $R^2(\tau)$, que representa a proporção da variância observada na variável dependente que é explicada pela variância na variável independente:

$$R^2(\tau) = \text{Var}(\hat{y}_{ij}) / \text{Var}(y_{ij}),$$

em que: $\text{Var}(\hat{y}_{ij})$ é a variância nos valores estimados e $\text{Var}(y_{ij})$ é a variância nos valores observados.

Conforme Nascimento et al. (2010), pode-se classificar os genótipos de duas maneiras: com estabilidade ou previsibilidade alta, quando $R^2(\tau)$ é maior que 70%; e com estabilidade ou previsibilidade baixa, quando $R^2(\tau)$ é menor que 70%. Para a estimação dos coeficientes de determinação, $\hat{\beta}_{0i}(\tau)$ e $\hat{\beta}_{1i}(\tau)$, da RQ, é necessário que se minimize antes a soma dos resíduos ponderados $y_i - \hat{y}_i$, de maneira que resíduos com valores positivos recebam peso τ e os com valores negativos, peso $1 - \tau$. Para tanto, devem-se obter os estimadores $\hat{\beta}_{0i}(\tau)$ e $\hat{\beta}_{1i}(\tau)$ que minimizem a seguinte equação:

$$\sum_{i=1}^n d_{\tau}(y_i, \hat{y}_i) = \tau \sum_{y_i \geq \beta_{0i}(\tau) + X_i \beta_{1i}(\tau)} |y_i - \beta_{0i}(\tau) - X_i \beta_{1i}(\tau)| + (1 - \tau) \dots \sum_{y_i < \beta_{0i}(\tau) + X_i \beta_{1i}(\tau)} |y_i - \beta_{0i}(\tau) - X_i \beta_{1i}(\tau)|,$$

em que d_τ é a distância entre y_i e \hat{y}_i .

A minimização dessa equação permite estimar os coeficientes da regressão, $\hat{\beta}_0(\tau)$ e $\hat{\beta}_1(\tau)$, o que é feito com base em algoritmos de programação linear, como descrito em Hao & Naiman (2007).

Após a estimação dos parâmetros, as hipóteses de interesse, $H_0: \beta_{li(\tau)} = 1$ e $H_1: \beta_{li(\tau)} \neq 1$, são avaliadas pelo teste t, dado por

$$t = \hat{\beta}_{1i}(\tau) - 1 / \sqrt{\hat{V}(\hat{\beta}_{1i}(\tau))}.$$

Essa estatística depende do número de graus de liberdade do resíduo na análise de variância conjunta e do nível de significância α . Em estudos de adaptabilidade e estabilidade, a variância de $\beta_1(\tau)$ é dada por $\sigma_c^2 = \text{QMR} / r$, em que r é o número de repetições que deram origem às médias submetidas à análise de adaptabilidade e estabilidade, e $\text{QMR} = \bar{\sigma}^2$ é o quadrado médio do resíduo da análise de variância conjunta.

Para avaliar a metodologia proposta, foram simulados 100 genótipos, considerando-se o seguinte modelo estatístico: $y_{ij} = \beta_{0i} + \beta_{1i} I_j + \varepsilon_{ij}$, em que: y_{ij} é a média do genótipo i no ambiente j ; β_{0i} é a constante da regressão referente ao i -ésimo genótipo; β_{1i} é o coeficiente de regressão, que mede a resposta do i -ésimo genótipo à variação do ambiente; I_j é o índice ambiental; e ε_i são os erros aleatórios gerados independentemente e identicamente distribuídos (i.i.d.), de acordo com a distribuição de probabilidade de interesse.

Para a obtenção dos valores de produtividade, os valores dos parâmetros β_0 , σ^2 e I_j foram obtidos do conjunto de dados avaliados por Nascimento et al. (2011). O valor de β_1 foi considerado igual à unidade, ou seja, genótipos com ampla adaptabilidade.

Foram avaliadas seis situações distintas, caracterizadas especificamente pela simulação de valores fenotípicos (y_i), com distribuições: simétrica; simétrica com “outliers”; assimétrica à direita; assimétrica à direita com “outliers”; assimétrica à esquerda; e assimétrica à esquerda com “outliers”.

A inserção de assimetria à direita ou à esquerda foi realizada, respetivamente, por meio da adição ou da subtração aos valores fenotípicos de resíduos amostrados de uma distribuição exponencial. No caso de distribuição simétrica dos valores fenotípicos, foram simulados resíduos independentes e identicamente distribuídos de acordo com distribuição normal com média zero e variância σ_e^2 , ou seja, $e_i \sim N(0, \sigma_e^2)$.

Para a simulação de “outliers”, considerou-se, como ponto de corte, o valor de $2/n^{0,5} = 2/20^{0,5} = 0,44$, para a medida de influência DFBeta (Belsley et al., 1980), ou seja: o conjunto de observações simuladas que apresentasse valor de DFBeta superior a 0,44 foi considerado como “outlier” (Montgomery et al., 2012). Para a inserção desses valores fenotípicos discrepantes no conjunto de dados, considerou-se o valor equivalente a 1,5 vezes a amplitude do interquartil, acrescido de uma constante (Montgomery & Runger, 2009).

Para avaliar o desempenho da regressão quantílica na estimação dos coeficientes, foram calculados os valores dos erros quadráticos médios (EQM) associados a cada estimador. Os EQM foram obtidos pelas seguintes equações:

$$\text{EQM}(\beta_{1i})_i = \frac{1}{100} \sum_{k=1}^{100} (\hat{\beta}_{1ik} - 1)^2,$$

em que $\hat{\beta}_{1ik}$ é o coeficiente de regressão estimado do genótipo i , na k -ésima repetição; e

$$\text{EQM}(\beta_{0i})_i = \frac{1}{100} \sum_{k=1}^{100} (\hat{\beta}_{0ik} - 1.176)^2,$$

em que $\hat{\beta}_{0ik}$ é a constante da regressão estimada para o genótipo i , na k -ésima repetição.

Os valores das estimativas obtidos com a metodologia proposta foram comparados aos obtidos com as de Eberhart & Russell (1966) e de regressão não paramétrica (Nascimento et al., 2010). A comparação dessas metodologias é interessante, uma vez que a RQ, à semelhança da metodologia de Eberhart & Russell (1966), é baseada em regressão e, assim como a metodologia proposta por Nascimento et al. (2010), é indicada na presença de pontos extremos, no conjunto de observações avaliadas.

O parâmetro de adaptabilidade foi estimado por meio de funções quantílicas, com valores de quantil estimados (τ) em 0,25, 0,50 e 0,75. Vale ressaltar que os valores de τ foram escolhidos para obtenção de uma melhor explicação para a relação entre variação ambiental e resposta fenotípica. Dessa forma, em distribuições de valores fenotípicos assimétricos à direita e à esquerda, espera-se que os quantis 0,25 e 0,75 representem melhor a informação de interesse. Já o valor de $\tau = 0,50$ representa a regressão mediana, que, no caso de distribuições simétricas, é equivalente à regressão na média.

Após a obtenção da estimativa do parâmetro de adaptabilidade em todas as metodologias, aplicou-se o teste t para verificar a significância da hipótese $H_0: \beta_{1i} = 1$. Posteriormente, foram calculadas as percentagens de falsos positivos, ou seja, a proporção em que a hipótese $\beta_1 = 1$ foi rejeitada, tendo-se considerado que os dados foram simulados de acordo com o modelo linear avaliado ($y_{ij} = \beta_{0i} + \beta_{1i} I_j + e_{ij}$).

Para a aplicação da metodologia proposta a um conjunto de dados reais, foram utilizados dados do experimento instalado em junho de 2004, no Campo Experimental da Embrapa Pecuária Sudeste, em São Carlos, SP. Utilizou-se o delineamento experimental de blocos ao acaso, com duas repetições, com as parcelas constituídas de cinco fileiras de 5 m de comprimento. Considerou-se uma fileira de cada lado, e os 0,50 m de cada extremidade da parcela como bordadura. Foi avaliada a produção de matéria seca de 92 acessos de alfafa. Os cortes foram feitos a 5 cm do solo, sempre que as cultivares apresentassem 10% de floração. Foram realizados 20 cortes, no período de novembro de 2004 a junho de 2006. Os cortes foram caracterizados como diferentes condições ambientais, visto que foram realizados em diferentes épocas.

Da mesma forma que no estudo de Nascimento et al. (2010), definiu-se como a medida da influência de um ponto, a variação em módulo entre os estimadores do coeficiente de inclinação, estimados pelos métodos de mínimos quadrados ($\hat{\beta}_{1i}^*$) e de regressão não paramétrica ($\hat{\beta}_{1i}^*$), para cada genótipo, isto é: $\Delta\hat{\beta}_{1i} = |\hat{\beta}_{1i}^* - \hat{\beta}_{1i}|$, em que um valor de 0,05 indica a existência de um ponto extremo.

Para verificar a assimetria dos dados, utilizou-se o teste de D'Agostino (1970), cuja hipótese de nulidade pressupõe que os dados amostrais são provenientes de uma população normalmente distribuída, ou seja, simétrica, com média μ e variância σ^2 .

Utilizou-se o programa livre R (R Core Team, 2014) para simulação e avaliação da adaptabilidade e da estabilidade pelo método proposto. O ajuste do modelo da RQ foi realizado com uso da função `rq` do pacote `quantreg` (Koenker, 2014). Para ajuste do modelo da regressão não paramétrica, utilizou-se a função `mblm` do pacote `mblm` (Komsta, 2013). A aplicação do teste de D'Agostino (1970) foi feita com o pacote `moments` (Komsta, 2012).

As rotinas computacionais implementadas no programa R estão presentes no site do Departamento de Estatística da Universidade Federal de Viçosa (2014).

Resultados e Discussão

De acordo com os valores de EQM (β_1), na ausência de “outliers”, a RQ ($\tau = 0,50$) que considera a função quantílica mediana apresentou resultados semelhantes aos obtidos com a metodologia de Eberhart & Russell (1966), em todas as situações avaliadas (Tabela 1). No caso de distribuições assimétricas à direita e à esquerda, verificaram-se menores valores de EQM [EQM (β_0) e EQM (β_1)] e de falsos positivos quando foram utilizadas as funções quantílicas RQ ($\tau = 0,25$) e RQ ($\tau = 0,75$), respectivamente. Esses resultados são condizentes com o fato de que a RQ permite generalizar o ajuste da relação funcional avaliada, para um quantil que represente essa relação de maneira mais adequada.

Na presença de “outliers” (Tabela 2), observaram-se melhores resultados (EQM e falsos positivos) para a RQ ($\tau = 0,50$) e para a regressão não paramétrica, se a distribuição era simétrica. Para distribuições assimétricas, houve redução tanto nos valores de EQM (β_1) quanto no de falsos positivos quando foram utilizadas as funções quantílicas RQ ($\tau = 0,25$) e RQ

Tabela 1. Erros quadráticos médios (EQM) e percentagem de falsos positivos obtidos com as metodologias de Eberhart & Russell (1966), regressão quantílica (RQ) e regressão não paramétrica, na ausência de “outliers”.

Método	EQM (β_1)	EQM (β_0)	Falso positivo
Distribuição assimétrica à direita			
Eberhart & Russell	0,0043	29.121	5,14
RQ ($\tau = 0,25$)	0,0019	3.593	1,09
RQ ($\tau = 0,50$)	0,0044	16.698	5,13
RQ ($\tau = 0,75$)	0,0125	61.592	21,90
Regressão não paramétrica	0,0048	17.011	6,63
Distribuição assimétrica à esquerda			
Eberhart & Russell	0,0043	145.409	5,31
RQ ($\tau = 0,25$)	0,0125	200.795	21,80
RQ ($\tau = 0,50$)	0,0044	122.715	5,49
RQ ($\tau = 0,75$)	0,0019	88.892	1,25
Regressão não paramétrica	0,0048	123.985	6,72
Distribuição simétrica			
Eberhart & Russell	0,0043	1.384.412	2,35
RQ ($\tau = 0,25$)	0,0082	1.136.971	7,85
RQ ($\tau = 0,50$)	0,0069	1.385.132	5,77
RQ ($\tau = 0,75$)	0,0080	1.658.201	7,13
Regressão não paramétrica	0,0087	1.384.246	15,70

($\tau = 0,75$), para distribuições assimétricas à direita e à esquerda, respectivamente.

No caso de distribuição assimétrica, a regressão não paramétrica obteve resultados inferiores à RQ (Tabela 2). Esse resultado evidencia que, apesar da regressão não paramétrica (Nascimento et al., 2010) ser menos influenciada por pontos extremos, ela não é tão eficaz para explicar relações funcionais em que os valores fenotípicos são assimétricos. Esse resultado decorre da utilização, na RQ, da moda para caracterizar conjuntos de valores com distribuição assimétrica; no presente trabalho, a moda foi aproximada pelos quantis $\tau = 0,75$ e $\tau = 0,25$, para distribuições assimétricas à esquerda e à direita, respectivamente.

De modo geral, independentemente do tipo de assimetria observada nos valores fenotípicos, ou da presença ou da ausência de “outliers”, a RQ possibilitou resultados similares ou melhores do que os obtidos com os demais métodos (Tabelas 1 e 2). Assim, no caso de valores fenotípicos simétricos, deve-se averiguar a presença de “outliers”. Se estes estiverem presentes, deve-se utilizar a RQ ($\tau = 0,50$); se não, tanto a metodologia de Eberhart & Russell (1966) quanto a RQ ($\tau = 0,50$) são adequadas. No caso

de distribuições assimétricas, sugere-se a RQ com $\tau = 0,25$, para assimetria à direita, e a RQ com $\tau = 0,75$, para assimetria à esquerda, independentemente da presença de “outliers”.

No experimento com alfafa, observou-se interação genótipos x ambientes ($p < 0,01$), que evidencia a necessidade de estudo pormenorizado da resposta das cultivares frente a variações ambientais (Tabela 3).

A Tabela 4 apresenta os 34 genótipos que obtiveram classificações de adaptabilidade discordantes de acordo com as metodologias avaliadas. A Tabela completa, com a classificação dos 92 genótipos, encontra-se no site do Departamento de Estatística da Universidade Federal de Viçosa (2014). Desses 34 genótipos, 29 apresentaram distribuições de valores fenotípicos simétricos, e 5, de valores assimétricos à direita, pelo teste de D’Agostino (1970). Além disso, 24 genótipos apresentaram $\Delta\hat{\beta}_{li} \geq 0,05$, ou seja, apresentavam “outliers”. Portanto, nesses casos, deve-se utilizar a RQ, conforme as considerações feitas anteriormente.

Os genótipos Primavera, Eterna, DK 177, Lujan, DK 194, P 5715, Rocio e GT 13 R Plus, que apresentam valores fenotípicos simétricos com presença de “outliers”, foram classificados como de adaptabilidade geral, tanto pela regressão não paramétrica quanto pela RQ ($\tau = 0,50$). Já os genótipos LE Semit 711 e Hunterfield, classificados como de adaptabilidade geral pela RQ ($\tau = 0,50$), foram considerados de adaptabilidade específica a ambientes favoráveis pela regressão não paramétrica. Esse resultado é indicativo de que a influência do “outlier” na estimação do parâmetro de adaptabilidade é maior na regressão não

Tabela 2. Erros quadráticos médios (EQM) e percentagem de falsos positivos obtidos com as metodologias de Eberhart & Russell (1966), regressão quantílica (RQ) e regressão não paramétrica, com presença de outliers.

Método	EQM (β_1)	EQM (β_0)	Falso positivo
Distribuição assimétrica à direita, com outliers			
Eberhart & Russell	0,1848	64.677	100
RQ ($\tau = 0,25$)	0,0022	3.859	2
RQ ($\tau = 0,50$)	0,0069	18.144	11
RQ ($\tau = 0,75$)	0,0974	99.255	70
Regressão não paramétrica	0,0060	15.514	10
Distribuição assimétrica à esquerda, com outliers			
Eberhart & Russell	0,1821	4.747.753	100
RQ ($\tau = 0,25$)	0,1238	4.418.813	67
RQ ($\tau = 0,50$)	0,0077	5.316.426	15
RQ ($\tau = 0,75$)	0,0023	5.668.763	4
Regressão não paramétrica	0,0064	5.345.482	10
Distribuição simétrica, com outliers			
Eberhart & Russell	0,0855	1.529.605	100
RQ ($\tau = 0,25$)	0,0169	1.095.687	40
RQ ($\tau = 0,50$)	0,0128	1.350.122	21
RQ ($\tau = 0,75$)	0,05635	1.788.108	61
Regressão não paramétrica	0,0202	1.326.998	39

Tabela 3. Análise de variância conjunta e média geral da produção de matéria seca de 92 genótipos de alfafa (*Medicago sativa*), em 20 cortes, em experimento conduzido de novembro de 2004 a junho de 2006.

Fonte de variação	GL	Quadrado médio
Blocos	1	2.002.415
Cultivar	91	1.384.476*
Erro a	91	574.270
Corte	91	62.331.023*
Erro b	91	946.918
Interação cultivar x corte	1.729	60.682**
Erro c	1.729	55.851
Média (kg ha ⁻¹)	1.176	

*e **Significativo pelo teste F, a 5 e 1% de probabilidade, respectivamente. Fonte: Nascimento et al. (2011).

paramétrica que na RQ ($\tau = 0,50$). Resultados inversos foram encontrados para os genótipos Aca 900, Medina, DK 193 e Rio Grande, que foram classificados como de adaptabilidade geral pelo método da RQ ($\tau = 0,50$), e de adaptabilidade específica a ambientes desfavoráveis pela regressão não paramétrica.

Tabela 4. Classificação dos genótipos quanto à adaptabilidade e à estabilidade, no caso de classificações discordantes, pelas três metodologias avaliadas.

Genótipo	Eberhart & Russell (1966)	Regressão quantílica		Regressão não paramétrica	$\Delta\beta_{ii}$
		($\tau=0,25$)	($\tau=0,50$)		
Primavera ^{ns}	D. P.	D. P.	G. P.	G. P.	0,105
LE Semit 711 ^{ns}	G. P.	F. P.	G. P.	F. P.	0,142
Topper ^{ns}	D. I.	G. P.	D. P.	G. P.	0,198
LE N 3 ^{ns}	G. P.	D. P.	G. P.	G. P.	0,017
Sequel HR ^{ns}	G. P.	G. P.	G. P.	F. P.	0,045
Victoria ^{ns(1)}	G. P.	D. I.	G. P.	G. P.	0,025
Aca 900 ^{ns}	G. P.	G. P.	G. P.	D. P.	0,135
5 939 ^{ns}	G. I.	G. P.	F. P.	G. I.	0,136
Medina ^{ns}	G. P.	G. P.	G. P.	D. P.	0,093
N 910*	G. P.	G. P.	G. P.	D. P.	0,182
Coronado*	G. P.	D. P.	D. P.	D. P.	0,084
Eterna ^{ns}	G. P.	D. P.	G. P.	G. P.	0,057
DK 193 ^{ns}	G. I.	G. P.	G. P.	D. I.	0,092
Crioula ^{ns}	G. P.	G. P.	F. P.	G. P.	0,050
DK 181 ^{ns}	G. P.	G. P.	F. P.	G. P.	0,005
Activa*	F. P.	G. I.	F. P.	F. P.	0,118
Aurora*	F. P.	G. I.	G. P.	G. I.	0,177
Hunterfield ^{ns}	G. P.	G. P.	G. P.	F. P.	0,251
P 105 ^{ns}	G. P.	F. P.	G. P.	G. P.	0,007
WL 516*	D. I.	G. I.	G. I.	G. I.	0,317
Tahoe ^{ns}	G. P.	G. P.	F. P.	F. P.	0,033
DK 167 ^{ns}	G. P.	G. P.	F. P.	G. P.	0,054
DK 177 ^{ns}	D. P.	G. P.	G. P.	G. P.	0,151
Maxidor ^{ns}	G. P.	G. P.	F. P.	G. P.	0,026
Tango ^{ns}	D. P.	D. P.	D. P.	G. P.	0,033
Rio Grande ^{ns}	G. P.	G. P.	G. P.	D. I.	0,165
Key II ^{ns}	F. P.	F. P.	F. P.	G. I.	0,191
Lujan ^{ns}	D. P.	G. P.	G. P.	G. P.	0,343
DK 194 ^{ns}	G. P.	G. P.	D. P.	D. P.	0,082
P 5715 ^{ns}	G. P.	D. P.	G. P.	G. P.	0,051
Aca 901 ^{ns}	F. P.	G. P.	F. P.	F. P.	0,026
Rocio ^{ns}	D. P.	G. P.	G. P.	G. P.	0,088
GT 13 R Plus ^{ns}	G. P.	G. P.	F. P.	F. P.	0,051
DK 187 R ^{ns}	G. I.	D. P.	G. P.	D. P.	0,026

*Significativo pelo teste de D'Agostino, a 5% de probabilidade. D, ambientes desfavoráveis; F, ambientes favoráveis; G, ambientes com adaptabilidade geral; I, imprevisível; e P, previsível. ^{ns}Não significativo pelo teste de D'Agostino, a 5% de probabilidade. $\Delta\beta_{ii} = |\beta^{ER}_{ii} - \beta^{NP}_{ii}|$, em que ER e NP referem-se às metodologias de Eberhart & Russell e não paramétrica, respectivamente. ⁽¹⁾Victoria SP INTA.

Os genótipos LE N 3, Sequel HR, Victoria SP INTA, P 105, Tango, Aca 901 e DK 187 R, com distribuição simétrica dos valores fenotípicos sem “outliers”, apresentaram a mesma classificação (adaptabilidade geral) nas metodologias de Eberhart & Russell (1966) e RQ ($\tau = 0,50$). Os genótipos Crioula, DK 181, Tahoe e Maxidor foram classificados como de adaptabilidade geral, quando avaliados pelo método de Eberhart & Russell (1966), e de adaptabilidade específica a ambientes favoráveis, quando avaliados pela RQ ($\tau = 0,50$). Nessa situação, deve-se utilizar a classificação dada por Eberhart & Russell (1966), uma vez que essa metodologia apresentou menor taxa de falsos positivos, de acordo com os dados simulados (Tabela 1).

Os genótipos Coronado, Activa, Aurora e WL 516, classificados como assimétricos pelo teste de D'Agostino (1970), e discrepantes quanto às classificações de adaptabilidade e estabilidade fenotípica (Tabela 4), apresentaram estimativas dos parâmetros de adaptabilidade mais adequadas com o uso da RQ ($\tau = 0,25$). O genótipo WL 516 foi classificado como de adaptabilidade específica a ambientes desfavoráveis, na metodologia de Eberhart & Russell (1966) e no trabalho de Ferreira et al. (2004); entretanto, no presente trabalho, este genótipo foi classificado como de adaptabilidade geral pela RQ ($\tau = 0,25$).

O genótipo P 5715 foi classificado como de adaptabilidade geral, tanto pela metodologia proposta no presente trabalho, com RQ ($\tau = 0,50$), quanto na metodologia utilizada por Vasconcelos et al. (2008).

Utilizou-se $R^2 = \text{Var}(\hat{y}_{ij})/\text{Var}(y_{ij})$, para avaliação da regressão não paramétrica e da RQ, já que, em razão da falta de ortogonalidade no modelo, a soma de quadrados total não pode ser decomposta em soma do quadrado da regressão mais soma do quadrado do resíduo, o que acarretaria estimativas negativas para o primeiro parâmetro.

Entre os fenótipos que apresentaram distribuição assimétrica nos valores fenotípicos, Activa e Aurora foram classificados como previsíveis, pelo método de Eberhart & Russell (1966). No entanto, quando analisados pela RQ com $\tau = 0,25$, esses genótipos foram classificados como imprevisíveis. Os três genótipos restantes obtiveram a mesma classificação nas duas metodologias: Coronado e N 910 foram classificados como previsíveis, e WL 516 como imprevisível.

Conclusões

1. O método proposto apresenta resultados iguais ou superiores às metodologias tradicionais avaliadas.

2. Para que se possa avaliar a interação genótipo x ambiente de forma eficiente, é importante que se avalie a simetria na distribuição dos valores fenotípicos, e a metodologia de regressão quantílica (RQ) deve ser usada no caso de distribuições assimétricas.

3. Para distribuições simétricas, na presença de “outliers”, deve-se utilizar a RQ com valor de quantil (τ) de 0,50; na ausência de “outliers”, pode-se utilizar tanto a metodologia de Eberhart & Russell quanto a RQ ($\tau = 0,50$).

4. Para distribuições assimétricas, indica-se o uso da RQ com $\tau = 0,25$, para assimetria à direita, e com $\tau = 0,75$, para assimetria à esquerda, independentemente da presença de “outliers”.

Agradecimentos

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) e à Fundação Arthur Bernardes (Funarbe), pela concessão de bolsas.

Referências

- BELSLEY, D.A.; KUH, E.; WELSCH, R.E. **Regression diagnostics**: identifying influential observations and sources of collinearity. London: Chapman & Hall, 1980. 292p.
- BEYERLEIN, A.; VON KRIES, R.; NESS, A.R.; ONG, K.K. Genetic markers of obesity risk: stronger associations with body composition in overweight compared to normal-weight children. **PLoS ONE**, v.6, 2011. DOI: 10.1371/journal.pone.0019057.
- CRUZ, C.D.; TORRES, R.A.; VENCOVSKY, R. An alternative approach to the stability analysis proposed by Silva and Barreto. **Revista Brasileira de Genética**, v.12, p.567-580, 1989.
- D'AGOSTINO, R.B. Transformation to normality of the null distribution of G1. **Biometrika**, v.57, p.679-681, 1970. DOI: 10.1093/biomet/57.3.679.
- EBERHART, S.A.; RUSSELL, W.A. Stability parameters for comparing varieties. **Crop Science**, v.6, p.36-40, 1966. DOI: 10.2135/cropsci1966.0011183X000600010011x.
- FERREIRA, R. de P.; BOTREL, M. de A.; RUGGIERI, A.C.; PEREIRA, A.V.; COELHO, A.D.F.; LÉDO, F.J. da S.; CRUZ, C.D. Adaptabilidade e estabilidade de cultivares de alfafa em relação a diferentes épocas de corte. **Ciência Rural**, v.34, p.265-269, 2004. DOI: 10.1590/S0103-84782004000100041.
- GAUCH JR, H.G. Statistical analysis of yield trials by AMMI and GGE. **Crop Science**, v.46, p.1488-1500, 2006. DOI: 10.2135/cropsci2005.07-0193.
- HAO, L.; NAIMAN, D.Q. **Quantile regression**. London: SAGE, 2007. 136p. (Quantitative applications in the social sciences, 149).
- KOENKER, R. **quantreg**: quantile regression. 2014. Available at: <<http://CRAN.R-project.org/package=quantreg>>. Accessed on: 12 Jul. 2014.
- KOMSTA, L. **moments**: moments, cumulants, skewness, kurtosis and related tests. 2012. Available at: <<http://cran.r-project.org/web/packages/moments/index.html>>. Accessed on: 12 Jul. 2014.
- KOMSTA, L. **mblm**: median-based linear models. 2013. Available at: <<http://cran.r-project.org/web/packages/mblm/index.html>>. Accessed on: 12 Jul. 2014.
- LIN, C.S.; BINNS, M.R. A superiority measure of cultivar performance for cultivar x location data. **Canadian Journal of Plant Science**, v.68, p.193-198, 1988. DOI: 10.4141/cjps88-018.
- MONTGOMERY, D.C.; PECK, E.A.; VINING, G.G. **Introduction to linear regression analysis**. 5th ed. Hoboken: John Wiley & Sons, 2012. 645p. (Wiley series in probability and statistics, 821).
- MONTGOMERY, D.C.; RUNGER, G.C. **Estatística aplicada e probabilidade para Engenheiros**. 4.ed. Rio de Janeiro: LTC, 2009. 493p.
- NASCIMENTO, A.C.C.; LIMA, J.E. de; BRAGA, M.J.; NASCIMENTO, M.; GOMES, A.P. Eficiência técnica da atividade leiteira em Minas Gerais: uma aplicação de regressão quantílica. **Revista Brasileira de Zootecnia**, v.41, p.783-789, 2012. DOI: 10.1590/S1516-35982012000300043.
- NASCIMENTO, M.; SILVA, F.F. e; SÁFADI, T.; NASCIMENTO, A.C.C.; FERREIRA, R. de P.; CRUZ, C.D. Abordagem bayesiana para avaliação da adaptabilidade e estabilidade de genótipos de alfafa. **Pesquisa Agropecuária Brasileira**, v.46, p.26-32, 2011. DOI: 10.1590/S0100-204X2011000100004.
- NASCIMENTO, M.; CRUZ, C.D.; CAMPANA, A.C.M.; TOMAZ, R.S.; SALGADO, C.C.; FERREIRA, R. de P. Alteração no método centroide de avaliação da adaptabilidade genotípica. **Pesquisa Agropecuária Brasileira**, v.44, p.263-269, 2009a. DOI: 10.1590/S0100-204X2009000300007.
- NASCIMENTO, M.; FERREIRA, A.; CAMPANA, A.C.M.; SALGADO, C.C.; CRUZ, C.D. Multiple centroid methodology to analyze genotype adaptability. **Crop Breeding and Applied Biotechnology**, v.9, p.8-16, 2009b. DOI: 10.12702/1984-7033.v09n01a02.
- NASCIMENTO, M.; FERREIRA, A.; FERRÃO, R.G.; CAMPANA, A.C.M.; BHERING, L.L.; CRUZ, C.D.; FERRÃO, M.A.G.; FONSECA, A.F.A. da. Adaptabilidade e estabilidade via regressão não paramétrica em genótipos de café. **Pesquisa Agropecuária Brasileira**, v.45, p.41-48, 2010. DOI: 10.1590/S0100-204X2010000100006.
- R CORE TEAM. **R**: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2014. Available at: <<http://www.r-project.org/>>. Accessed on: 12 Jul. 2014.
- SILVA, E.N. da; PORTO JÚNIOR, S. da S. Sistema financeiro e crescimento econômico: uma aplicação de regressão quantílica. **Economia Aplicada**, v.10, p.425-442, 2006. DOI: 10.1590/S1413-80502006000300007.

UNIVERSIDADE FEDERAL DE VIÇOSA. Departamento de Estatística. **Materiais disponíveis**. Disponível em: <<http://www.det.ufv.br/~moyses/links.php>>. Acesso em: 12 jul. 2014.

VASCONCELOS, E.S. de; BARIONI JÚNIOR, W.; CRUZ, C.D.; FERREIRA, R. de P.; RASSINI, J.B.; VILELA, D. Seleção de genótipos de alfafa pela adaptabilidade e

estabilidade da produção de matéria seca. **Acta Scientiarum. Agronomy**, v.30, p.339-343, 2008. DOI: 10.4025/actasciagron.v30i3.3511.

YAN, W.; TINKER, N.A. Biplot analysis of multi-environment trial data: principles and applications. **Canadian Journal of Plant Science**, v.86, p.623-645, 2006. DOI: 10.4141/P05-169.

Recebido em 7 de agosto de 2014 e aprovado em 23 de março de 2015