

MODELOS DE REGRESSÃO EM COMPONENTES PRINCIPAIS PARA A PREVISÃO DE PRODUTIVIDADES DE CAFEIROS NO SUL DO ESTADO DE MINAS GERAIS¹

Luiz Gonsaga de CARVALHO², Gilberto C. SEDIYAMA³, Paulo Roberto CECON⁴, Helena Maria Ramos ALVES⁵

1. INTRODUÇÃO

A previsão de safras agrícolas no Brasil tem sido um grande problema para o agronegócio. Procurando elaborar modelos de previsão, diversos trabalhos existentes na literatura mostram a importância das condições climáticas na produção agrícola. Na cultura do café, um fator importante que interfere na variação de sua produção, próprio de sua natureza fisiológica, é a alternância bienal, com safras altas e baixas, RENA & MAESTRI (1985).

DOORENBOS & KASSAM (1979) propõem o modelo desenvolvido por Stewart et al. como alternativa de modelagem tendo a produtividade como função de parâmetros hídricos, sendo este modelo trabalhado por PICINI (1998) e adaptado para a cultura do café acrescentando a influência da produtividade do ano anterior.

Uma técnica bastante útil em modelagem é a análise de componentes principais, os quais são independentes entre si e estimados com o propósito de reter, em ordem decrescente, o máximo de informação, em termos de variação total, contida nos dados iniciais, permitindo, com isso, reduzir a dimensão do conjunto original. Desta forma, pode-se quando se tem o número de observações inferior à quantidade de variáveis, realizar a regressão linear com os componentes principais que retêm a maior parte da variância original das variáveis, permitindo com que o número de parâmetros a serem estimados sejam menor do que o número de observações.

Assim, o objetivo deste trabalho foi parametrizar um modelo de regressão linear múltipla em componentes principais para a previsão de produtividade da cultura do café utilizando um conjunto de variáveis além dos componentes hídricos, para três localidades da região Sul do Estado de Minas Gerais.

2. MATERIAL E MÉTODOS

Os dados de produtividades de café, utilizados são representativos dos municípios de Alfenas (lat.: 21° 20' S, long.: 45° 56' W, alt.: 844 m), Monte Belo (lat.: 21° 23' S, long.: 46° 14' W, alt.: 940 m) e São Sebastião do Paraíso (lat.: 20° 54' S, long.: 47° 03' W, alt.: 982 m), com caracterização da bienalidade de produção bem definidas. As produtividades potenciais adotadas foram 99, 81 e 88 sacas ha⁻¹ para os três municípios, respectivamente.

As estimativas da disponibilidade hídrica do solo, para uma capacidade de água disponível de 100 mm, foram realizadas pelo método proposto por Thornthwaite & Mather. O valor do coeficiente da

cultura (Kc) utilizado para a estimativa da evapotranspiração da cultura foi igual a 1 (um), também utilizado por PICINI (1998). Para cada município os dados diários de temperatura e precipitação pluvial foram agrupados em escala mensal, obtendo-se as respectivas médias, efetuando-se em seguida o balanço hídrico seriado para o mesmo período das produtividades observadas.

Posteriormente, os dados e componentes do balanço hídrico, foram separados por uma seqüência de períodos trimestrais, de acordo com MATELLO (1991), assim definida: jul/ago/set; out/nov/dez; jan/fev/mar e abr/maio/jun. Para cada um destes períodos e para todo o período das produtividades, obteve-se as médias dos componentes do balanço hídrico e demais variáveis relacionadas à temperatura, totalizando 20 variáveis, as quais foram submetidas ao modelo de regressão. Para cada local representado por sua respectiva gleba, os modelos consistiram em submeter à análise de regressão linear múltipla, a produtividade da cultura do café como função deste conjunto de variáveis. Como a quantidade destas variáveis foi superior ao número de observações, adotou-se a técnica de componentes principais para transformar estas variáveis em um novo conjunto, ou seja, os chamados escores dos componentes principais.

3. RESULTADOS E DISCUSSÃO

As variâncias acumuladas para os três primeiros componentes principais foram 84,04; 84,68 e 90,99% para Alfenas, Monte Belo e São Sebastião do Paraíso, respectivamente, sendo suficientes para representarem as variáveis analisadas.

Na Tabela 1, verifica-se que, para Alfenas, as estimativas dos parâmetros são significativas ($P \leq 0,0787$) indicando que as variáveis transformadas em componentes principais contribuíram de maneira satisfatória. Isto é reforçado pelo coeficiente de determinação ($R^2 = 0,88$). Para Monte Belo, a estimativa do parâmetro " β_3 " possui baixa significância, contudo manteve-se o componente principal correspondente (Z_3) na regressão linear por praticamente não alterar os resultados com a sua retirada e por coerência ao modelo anterior. Os demais parâmetros mostram-se com significâncias relevantes ($P \leq 0,1131$) e, o coeficiente de determinação ($R^2 = 0,75$) é também satisfatório. Em São Sebastião do Paraíso, a estimativa do parâmetro " β_1 " possui um nível de significância relativamente baixo ($P \leq 0,1889$), porém o componente principal

¹ Parte da tese de doutorado do primeiro autor. Projeto financiado pelo CNPQ, Proc. N. 474244/2001-6 (NV), mod. APQ

² Dr. Prof. Adjunto, Departamento de Engenharia, DEG, Universidade Federal de Lavras, CP 37, CEP 37200-000, Lavras, MG. E-mail: lgonsaga@ufla.br.

³ Professor Titular, DEA/UFV. E-mail: sedyama@ufv.br.

⁴ Professor Titular, DPI/UFV. E-mail: cecon@dpi.ufv.br.

⁵ Pesquisadora, EMBRAPA. E-mail: helena@ufla.br.

correspondente (Z_1) por ser o de maior importância, foi mantido no modelo. Assim também, o parâmetro " β_0 " não é significativo, porém foi mantido no modelo para ser semelhante aos modelos anteriores e pouco alterar o resultado final com sua retirada. Os demais parâmetros, por suas significâncias ($P \leq 0,0935$), indicam seus níveis de importância dentro do modelo. Por sua vez, o coeficiente de determinação ($R^2 = 0,66$) pode ser considerado satisfatório.

Quadro 1. Coeficientes de determinação, testes F, coeficientes da regressão e testes t dos modelos para os municípios do Sul de Minas Gerais

Coef.	Alfenas		Monte Belo		São Sebastião do Paraíso	
	Modelo	Valor-P ⁽¹⁾	Modelo	Valor-P ⁽¹⁾	Modelo	Valor-P ⁽¹⁾
R^2	0,8825		0,7513		0,6649	
Teste F	0,0004		0,0044		0,0097	
β_0	0,6798	0,0191	1,6594	0,1131	-0,1109	0,8551
β_1	0,1719	0,0787	-0,7688	0,0011	-0,2308	0,1889
β_2	-0,7805	0,0002	0,4439	0,0571	-0,6448	0,0035
β_3	0,7034	0,0044	-0,2659	0,4987	0,5447	0,0935

⁽¹⁾ Valor do nível de significância dos parâmetros da equação de regressão pelo teste t.

Da Tabela 1, por exemplo, o modelo para Alfenas fica assim representado:

$$\hat{Y}_e = Y_p[0,6798 + 0,1719Z_1 - 0,7805Z_2 + 0,7034Z_3]$$

em que, \hat{Y}_e é a produtividade estimada (sacas ha^{-1}); Y_p a produtividade potencial (sacas ha^{-1}); Z_1 , Z_2 e Z_3 são os componentes principais, sendo estes funções lineares das 20 variáveis.

A Figura 1 mostra o comportamento do modelo para o município de Alfenas ao comparar com produtividades independentes à sua elaboração. Nota-se a similaridade aos dados observados. O mesmo foi constatado para os outros municípios.

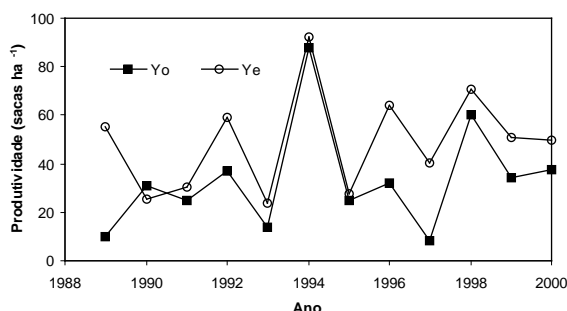


Figura 1. Produtividades observadas e estimadas pelo modelo para o município de Alfenas – MG.

Pela Figura 2, verifica-se o desempenho deste modelo apresentando $r^2 = 0,58$ e o índice de WILLMOTT et al. (1985), $d = 0,78$, indicando a performance da concordância entre valores observados e estimados. O índice "d", com variação entre 0 e 1, indica o grau de exatidão entre valores estimados e observados, sendo que quanto mais próximo de 1 (um) melhor a exatidão. Semelhantemente aos resultados de Alfenas, para

Monte Belo a regressão linear simples, com o ajuste da reta passando pela origem, apresentou $r^2 = 0,75$ indicando uma precisão relativamente satisfatória das estimativas. Por sua vez, o índice "d" ($d = 0,80$), mostrou boa concordância entre produtividades observadas e estimadas. Em São Sebastião do Paraíso, os resultados mostraram-se com maiores discrepâncias apresentando $r^2 = 0,15$ e $d = 0,58$.

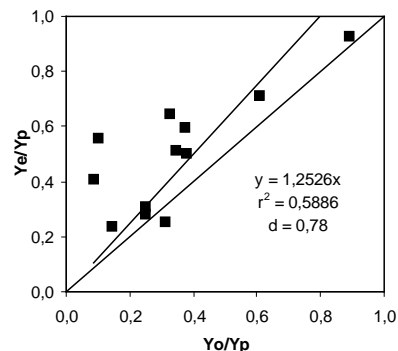


Figura 2. Desempenho entre produtividades observadas e estimadas pelo modelo para Alfenas – MG.

Observa-se ainda na Figuras 2, assim como para os demais municípios, que o modelo tende a superestimar a produtividade. Estas discrepâncias, também foram observadas por PICINI (1998) ao trabalhar com a cultura do café.

4. CONCLUSÃO

A tentativa de acrescentar novas variáveis ao modelo de Stewart et al. modificado por PICINI (1998), para prever produtividades de cafeeiros não se mostrou satisfatório, apresentando ainda grandes discrepâncias entre produtividades estimadas e observadas e, ocorrendo a tendência do modelo em superestimar as produtividades.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- DOORENBOS, J.; KASSAM, A.H. **Yield response to water**. Rome: FAO, 1979. 197p. (FAO Irrigation and Drainage Paper, 33).
- MATIELLO, J.B. **O café: do cultivo ao consumo**. São Paulo: Editora Globo, 1991. 319p. (Coleção do Agricultor - Grãos).
- PICINI, A.G. **Desenvolvimento e teste de modelos agrometeorológicos para a estimativa de produtividade do cafeeiro (*Coffea arabica* L.) a partir do monitoramento da disponibilidade hídrica do solo**. Piracicaba: ESALQ-USP, 1998. 132p. (Dissertação de Mestrado).
- RENA, A.B.; MAESTRI, M. Fisiologia do cafeeiro. **Informe Agropecuário**, Belo Horizonte, v.11, n.126, p.26-40, 1985.
- WILLMOTT, C.J.; ACKLESON, S.G.; DAVIS, R.E.; FEDDEMA, J.J.; KLINK, K.M.; LEGATES, D.R.; O'DONNELL, J.; ROWE, C.M. Statistics for the evaluation and comparison of models. **Journal of Geophysical Research**, Washington D.C., v.90, n.C5, p.8995-9005, 1985.