

APLICAÇÃO DA ANÁLISE HARMÔNICA POR SÉRIES DE FOURIER PARA A PREVISÃO DE PRODUTIVIDADE DA CULTURA DO CAFÉ NO ESTADO DE MINAS GERAIS¹

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

1. INTRODUÇÃO

A previsão de produtividade constitui uma ferramenta extremamente útil no sentido de permitir, ao agricultor e ao país, um melhor planejamento de das atividades agrícolas. Na cultura do café, a produtividade sofre oscilações sendo afetada por fatores climáticos e pelo efeito da bienalidade de produção, característica fisiológica própria dessa cultura conforme discute RENA & MAESTRI (1985).

Para o estudo de séries temporais que apresentam variações periódicas sazonais, a análise harmônica por séries de Fourier é aplicada para o desenvolvimento de modelos que permitem a estimativa da variável previsível em análise. Outra técnica bastante útil em modelagem é a análise multivariada de componentes principais, cujo objetivo é transformar um conjunto de variáveis, em um novo conjunto de variáveis não correlacionadas, ou seja, os componentes principais. 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, desde que sejam suficientes para representar as variáveis originais, reduzindo substancialmente as variáveis regressoras dentro do modelo linear de regressão múltipla, tendo-se assim que, o número de parâmetros a serem estimados passa a ser menor do que o número de observações.

Com base no exposto, este trabalho objetivou avaliar a aplicação da análise harmônica por séries de Fourier para a previsão de produtividades da cultura do café para o Estado de Minas Gerais.

2. MATERIAL E MÉTODOS

Os dados de produtividade da cultura do café foram obtidos em registros do IBGE (Instituto Brasileiro de Geografia e Estatística). Utilizou-se sete séries anuais de dados de produtividades médias correspondentes ao período de 1986-2000 (quinze anos) dos municípios de Alfenas, Guaxupé, Monte Belo, Lavras, São Sebastião do Paraíso, Varginha e Viçosa no Estado de Minas Gerais, cujas latitudes variam de 20° 45' S (Viçosa) a 21° 31' S (Varginha); longitudes de 42° 51' W (Viçosa) a 47° 03' W (São Sebastião do Paraíso) e altitudes de 690 m (Viçosa) a 982 m (São Sebastião do Paraíso). Dado o efeito bienal de produção do cafeeiro, cada série de dados de produtividades, foi submetida à análise harmônica por séries de Fourier de senos e cossenos conforme a expressão seguinte, SPIEGEL (1974).

$$f(x) = Y_0 + \sum_{n=1}^{\infty} (a_n \cos n\omega t_i + b_n \sin n\omega t_i) \quad (1)$$

em que, Y_0 = coeficiente das séries de Fourier (harmônico fundamental) - representa a própria média aritmética da série de dados observados; $n = 1, 2, \dots, \infty$, identifica o número de harmônicos; a_n e b_n = coeficientes de cosseno e seno, respectivamente, para o harmônico "n" das séries de Fourier; ω = frequência angular da onda da série de dados, sendo $\omega = 2\pi/T$ e T o período, sendo este correspondente ao número de dados observados; t_i = ordenação numérica dos dados correspondentes da série ($i = 0, 1, \dots, T-1$).

O número de harmônicos ajustados pelas séries de Fourier para cada conjunto de dados foi igual a 7 (sete), ou seja, $(n-1)/2$, pois as séries contêm 15 (quinze) observações. Foram utilizadas 33 variáveis predictoras, sendo uma referente à produtividade do ano anterior e 32 oriundas do balanço hídrico, para capacidade de água disponível no solo de 100 mm, e de elementos climáticos relacionadas à temperatura e exposição à radiação solar separados conforme as fases fenológicas de produção do cafeeiro de acordo com MATIELLO (1991), sendo assim definidas: Fase 1: período compreendido pelos meses de julho a setembro; Fase 2: de outubro a dezembro; Fase 3: janeiro a março e Fase 4: pelos meses de abril a junho. Como no presente estudo foram submetidas à análise de Fourier sete séries de dados de produtividades, representando sete municípios, obteve-se portanto, para cada coeficiente de ordem "n", seja a_n ou b_n , sete valores correspondentes. Como para cada um destes coeficientes com sete repetições (sete municípios) existe um conjunto de trinta e três variáveis, recorreu-se a técnica de componentes principais para transformar estas variáveis em um novo conjunto de variáveis, ou seja, os chamados escores dos componentes principais, permitindo assim efetuar as regressões. Os coeficientes obtidos (Y_0 , a_n e b_n) foram separados e submetidos como variáveis dependentes à análise de regressão linear múltipla em função dos escores dos componentes principais retidos. Voltando-se à equação 1, estes coeficientes foram substituídos por suas respectivas regressões, gerando o modelo prognóstico de produtividade de café para a região representada pelos municípios trabalhados.

3. RESULTADOS E DISCUSSÃO

Os três primeiros componentes principais do conjunto de variáveis utilizadas apresentaram a variância acumulada de 94,65%, sendo, portanto,

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

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

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

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

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

suficientes para representarem os dados originais com sucesso.

Verificou-se que as melhores regressões foram para os coeficientes Y_0 , a_3 e a_4 , cujos valores dos coeficientes de determinação (R^2) ficaram compreendidos entre 0,79 e 0,84; os níveis críticos dos testes F da regressão com $P \leq 0,151$ e os parâmetros da regressão (teste t) com $P \leq 0,586$. Para os demais coeficientes, as regressões não se apresentaram estatisticamente satisfatórias, mostrando não existir uma boa resposta de regressão dos coeficientes de Fourier nos componentes principais das variáveis de entrada do modelo. Contudo, a eliminação destas regressões não significativas praticamente não alterou o resultado final do modelo. Assim, para que houvesse uma similaridade no conjunto de parâmetros entre o modelo prognóstico e o diagnóstico (modelo devido ao ajuste das séries de Fourier aos dados de produtividades) resolveu-se manter todas as regressões no modelo final. Desta forma, o modelo originado, com base na equação 1, é representado pelo conjunto das seguintes equações:

$$\hat{Y}_R = Y_0 + \left[(a_1 \cos(1 * \frac{2\pi}{T} * t_i) + (b_1 \sin(1 * \frac{2\pi}{T} * t_i) \right] + \left[(a_2 \cos(2 * \frac{2\pi}{T} * t_i) + (b_2 \sin(2 * \frac{2\pi}{T} * t_i) \right] + \dots + \left[(a_7 \cos(7 * \frac{2\pi}{T} * t_i) + (b_7 \sin(7 * \frac{2\pi}{T} * t_i) \right] \quad (2)$$

em que, os coeficientes Y_0 , a_n e b_n ("n" variando de 1 a 7) são representados por:

$$Y_0 = 1489,58 + 19,91Z_1 - 75,73Z_2 - 72,78Z_3 \quad (3)$$

$$a_1 = 18,79 - 6,49Z_1 - 2,62Z_2 - 37,54Z_3 \quad (4)$$

$$b_7 = -40,38 - 14,12Z_1 + 9,87Z_2 - 4,66Z_3 \quad (5)$$

sendo Z_1 , Z_2 e Z_3 os componentes principais (funções lineares) das 33 variáveis de entrada. Na equação 2 o período (T) equivale à 15 (quinze) devido ser este modelo gerado de séries compostas por quinze anos de dados de produtividades e t_i o valor de ordem do dado. Como este modelo é de estimativa, t_i será sempre igual à 1 (um).

Realizado o teste do modelo, para o município de Três Pontas (lat.: 21° 22' S, long.: 45° 29' W, alt.: 900 m) verificaram-se que os erros relativos percentuais de comparação entre valores estimados e observados, variaram entre -39,54 e 85,58%. Apesar das discrepâncias, nota-se pela Figura 1, que o comportamento das produtividades estimadas tende a acompanhar as observadas. Por sua vez, a Figura 2, mostra o desempenho do modelo indicando um baixo valor para o índice de WILLMOTT et al. (1985) ($d = 0,48$) não apresentando portanto, boa concordância entre valores estimados e observados. Por outro lado, obtendo a regressão linear simples, com a reta de regressão passando pela origem, para as produtividades estimadas como função das observadas, obteve-se o coeficiente de determinação

muito baixo ($r^2 = 0,027$) indicando a baixa precisão do modelo em estimar a produtividade para a cultura do café e, verificando também, a tendência do modelo em subestimar a produtividade.

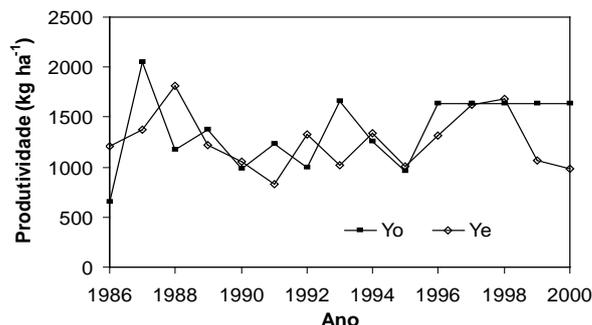


Figura 1. Produtividades observadas (Y_o) e estimadas (Y_e) pelo modelo para o município de Três Pontas - MG.

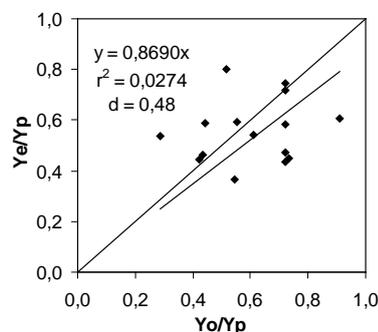


Figura 2. Produtividades relativas estimadas (Y_e/Y_p) e observadas (Y_o/Y_p) e respectivo desempenho estatístico do modelo para o município de Três Pontas - MG.

4. CONCLUSÃO

O modelo avaliado não se mostrou satisfatório na previsão de produtividade para a cultura do café, havendo a necessidade de se trabalhar com dados reais de produtividade representando diversos municípios, para melhor avaliar e testar seu desempenho.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- MATIELLO, J.B. **O café: do cultivo ao consumo**. São Paulo: Editora Globo, 1991. 319p. (Coleção do Agricultor - Grãos).
- SPIEGEL, M.R. **Fourier analysis**. New York: McGraw-Hill Book Company, 1974. 191p.
- 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.