

Modelagem de dados espectrais e meteorológicos na produtividade da cultura do café

Tiago Bernardes¹
Marcos Adami²
Maurício Alves Moreira³

¹ Centro Nacional de Monitoramento e Alerta de Desastres Naturais - CEMADEN
Rodovia Presidente Dutra, km 40 - 12630-000 – Cachoeira Paulista - SP, Brasil
tiago.bernardes@cemaden.gov.br

² Centro Regional da Amazônia - CRA
Parque de Ciência e Tecnologia do Guamá, 2651
CEP 66077-830 - Belém - PA, Brasil
marcos.adami@inpe.br

³ Instituto Nacional de Pesquisas Espaciais - INPE
Caixa Postal 515 - 12227-010 - São José dos Campos - SP, Brasil
mauricio@dsr.inpe.br

Abstract. Remote sensing and meteorological data were used to model coffee yield on the South of Minas Gerais state. Several vegetation indices derived parameters were applied to estimate an indicator factor of biennial bearing effect on coffee yield and to be incorporated into the agrometeorological model to estimate yield. The final model penalizes the potential yield as a function of: i) relative evapotranspiration (water penalization) obtained from meteorological data and ii) yield of previous year. Meteorological data were obtained from the regional weather prediction model ETA, and the yield of previous year was estimated from relationships between yield and remote sensing derived vegetation indices. The final model was evaluated at two levels: coffee plots and municipalities. Overall performance of the model at the level of coffee plots was considered inconclusive – errors varying from 13,73 to 17,46 ‘60-kg bags of green coffee per hectare’ (MSE) and 18,51 to 23,15 ‘60-kg bags of green coffee per hectare’ (RMSE), R^2 values varying from 0,50 to 0,69, d_1 index from 0,45 to 0,53 and d_2 index from 0,59 a 0,70. At the level of municipality the performance was considerably better – errors varying from 1.607 to 2.147 kg (MSE) and 2.490 to 3.526 kg (RMSE), R^2 values varying from 0,86 to 0,93, d_1 index from 0,78 to 0,83 and d_2 index from 0,92 to 0,96.

Palavras-chave: vegetation indices, agrometeorological model, coffee yield, índices de vegetação, modelo agrometeorológico, produtividade de café.

1. Introdução

No Brasil, a produtividade da cultura do café é estimada operacionalmente por métodos subjetivos ou em estudos científicos com algum nível de formalismo matemático como nos modelos agrometeorológicos (Piccini, et. al., 1999; Camargo et al., 2003; Carvalho et al., 2003; Santos et al., 2006). No segundo caso, dada a complexidade das relações envolvidas tais modelos operam, na maioria das vezes, de maneira pontual. No entanto, os modelos agrometeorológicos consideram predominantemente a influência de parâmetros meteorológicos na produtividade. Uma das características mais marcantes da cultura cafeeira é o fenômeno da bienalidade da produção, o qual proporciona baixas e altas produtividades em anos alternados. A produtividade é um fenômeno relacionado à área foliar das culturas, a qual pode ser estimada por meio de índices derivados de imagens orbitais (Funk e Budde, 2009; Brunsell et al., 2009; Bernardes et al., 2012). Assim, ela pode ser obtida espacialmente por meio da associação de dados de imagens de satélites aos modelos agrometeorológicos. O desenvolvimento de novas plataformas orbitais e sensores tem resultado numa grande variedade de produtos com melhorias expressivas em resoluções espacial, espectral, radiométrica e temporal. Tais avanços favorecem a descrição da variabilidade de padrões da

cultura cafeeira e obtenção de melhores correlações entre os dados espectrais e variáveis biofísicas da cultura, melhorando as possibilidades de aplicação de técnicas de sensoriamento remoto em estudos com café.

2. Metodologia

2.1. Área de estudo

Os estudos foram realizados na Mesoregião Sul/Sudoeste do Estado de Minas Gerais, onde a área cultivada com café corresponde mais da metade da área total de café no Estado. A região foi escolhida em função da sua representatividade no cenário nacional da produção de café, diversidade de ambientes e sistemas de cultivo e presença de instituições de excelência em pesquisa cafeeira potencialmente parceiras na geração e análise de dados. Três áreas piloto, representativas do núcleo cafeeiro na região, foram utilizadas para levantamentos de campo para coleta de dados. São elas: (i) Boa Esperança/Campo do Meio/Campos Gerais/Três Pontas; (ii) Alfenas e; (iii) São Sebastião do Paraíso/Monte Santo/Itamogi. Os dados levantados nestas áreas foram analisados individualmente para verificar ajuste dos modelos em escala local e também em conjunto para aplicar na escala municipal.

2.2. Dados

Foram utilizadas imagens TM/Landsat-5 como dados básicos para obter os índices de vegetação NDVI e o EVI. As imagens de anos anteriores também foram obtidas para identificar padrões de bienalidade da produção por meio dos índices de vegetação. Além dos índices de vegetação, foram extraídas das imagens as áreas cafeeiras visando aplicar as rotinas para cálculo da produtividade apenas dentro das áreas de café. Para realizar os cálculos de estimativa da produtividade do ano anterior a partir dos índices de vegetação foi utilizado um mapa base contendo as áreas cafeeiras, obtido a partir da interpretação de imagens do ano de 2007 (Moreira et al., 2007). Os procedimentos de correção radiométrica, correção atmosférica e normalização das imagens para geração dos índices de vegetação estão descritos detalhadamente em Bernardes et al. (2013) e Bernardes (2013).

Os dados meteorológicos utilizados no balanço hídrico foram derivados das previsões do modelo regional de previsão do tempo ETA: (i) temperatura média do ar ($^{\circ}\text{C}$), (ii) precipitação pluviométrica (mm.dia^{-1}), (iii) radiação solar ($\text{cal.cm}^{-2}.\text{dia}^{-1}$), (iv) velocidade do vento (km.dia^{-1}) e, (v) umidade relativa do ar (%). Os dados de altimetria, necessários para cálculo do fator W, utilizado para obter a evapotranspiração potencial, foram extraídos dos arquivos do ETA. Estes dados são fornecidos pelo Centro de Previsão de Tempo e Estudos Climáticos (CPTEC/INPE) em resoluções espaciais de 40 x 40 km e 20 x 20 km. Além da espacialização das variáveis em questão estes dados foram obtidos antecipadamente, melhorando as possibilidades de aplicar em modelos de previsão de produtividade.

Foram coletados em campo dados de produtividade das lavouras (variedades Catuaí e Mundo Novo) entre os anos de 2001 e 2010. Os dados foram divididos em dois conjuntos distribuídos espacialmente. Um para a construção de modelos empíricos e parametrização e outro para validar os resultados finais. Para avaliação do modelo em escala municipal foram utilizados também dados de produção total dos municípios disponíveis da base de dados do IBGE (IBGE, 2009).

2.3. Modelo

O aspecto chave da metodologia consiste em utilizar os dados de sensoriamento remoto (métricas de índices de vegetação) para gerar correlações com a produtividade de lavouras cafeeiras de modo a permitir a incorporação da produtividade do ano anterior (Yaa) ao modelo. Para tanto foram utilizados modelos de regressão entre índices de vegetação e

produtividade na estimativa da produtividade do ano anterior (Y_{aa}) a ser incorporada ao modelo agrometeorológico (Doorembos e Kassan, 1979) ajustado para a cultura cafeeira com adaptações para inclusão do efeito da produtividade do ano anterior conforme Santos e Camargo (2006) - Equação 01:

$$Y_{est} = Y_p \left\{ \left[k_{y_0} \left(\frac{Y_{aa}}{Y_p} \right) \right] + \left[\left[K_{y_i} \left(1 - \frac{ET_r}{ET_0} \right) \right] \right] \right\} \quad (01)$$

em que:

Y_{est} = produtividade estimada; k_{y_0} = coeficiente de penalização pela produtividade do ano anterior; Y_{aa} = produtividade do ano anterior (estimadas a partir de imagens de satélite); Y_p = produtividade potencial; k_{y_i} = coeficiente de resposta da cultura à disponibilidade hídrica no estágio i ; ET_r = Evapotranspiração real; ET_0 = Evapotranspiração potencial

Assim, foram obtidas regressões entre os índices de vegetação referentes aos talhões de café e a produtividade dos mesmos para estimativa da produtividade do ano anterior, dada a disposição de imagens para os anos anteriores. O procedimento de construção e seleção dos modelos de regressão empregados é descrito detalhadamente em Bernardes et al (2013) e Bernardes (2013).

2.4. Coeficientes de resposta (k_{y_0} e k_{y_i})

De posse das variáveis: Produtividade potencial (Y_p), Produtividade do ano anterior (Y_{aa}) e Componentes meteorológicas (ET_r/ET_0), necessárias ao modelo final definido pela Equação 01 foram obtidos os coeficientes de resposta (K_y) de cada componente.

Os coeficientes foram obtidos por regressão linear múltipla em que a Equação 01 é formulada com todos os parâmetros necessários em ambos os lados da igualdade – Produtividade observada (em lugar da produtividade estimada) e Produtividade potencial (Y_p), Produtividade do ano anterior (Y_{aa}) e Componentes meteorológicas (ET_r/ET_0). A regressão define então os coeficientes que otimizam a igualdade.

3. Resultados

3.1. Modelo de regressão para estimativa da produtividade do ano anterior Y_{aa}

A Tabela 01 apresenta três modelos para as variedades Catuaí, Mundo Novo e Catuaí+Mundo Novo, o número de unidades amostrais e de variáveis empregado em cada modelo, parâmetros estatísticos obtidos com dados de ambas as cultivares juntas (modelo geral) e dados das cultivares Catuaí e Mundo Novo isoladamente.

Tabela 01 - Parâmetros estatísticos obtidos para os modelos utilizando dados das cultivares Catuaí e Mundo Novo e dados agregados de ambas as variedades (Catuaí e Mundo Novo)

Modelo	Tamanho da amostra	Número de variáveis	Valor F	Pr > F	r	R ²
Catuaí (Cat)	47	4	26.60	<.0001	0,84	0,71
Mundo Novo (MN)	179	3	24.77	<.0001	0,53	0,29
Cat + MN	308	4	30.96	<.0001	0,52	0,28

A Equação 02 foi empregada para estimativa da produtividade do ano anterior a ser incluída no modelo final.

$$Y_{aa} = 9,1 - 31,65EVI^2 + 46,21SqEVI_{aa} - 174,28DifEV_{Imo} + 118,36DifSqEV_{Imo} \quad (02)$$

em que:

EVI² - EVI ao quadrado; SqEVI_{aa} - soma dos quadrados do EVI média do ano anterior; Dif2EVI_{mo} - diferença ao quadrado do EVI moda de dois anos consecutivos; DifSqEVI_{mo} - diferença da soma dos quadrados do EVI moda de 2 anos consecutivos;

3.2. Estimativas de produtividade com o modelo final

De posse de todas as variáveis necessárias e dos coeficientes de resposta (ky) para as componentes espectral (Produtividade do ano anterior – Y_{aa}) e meteorológicas (Evapotranspiração relativa nas diferentes fases fenológicas do cafeeiro – ETr/ET₀) foi obtido a partir das Equações 01 e 02 um modelo geral para estimativa da produtividade do café – Equação 03:

$$Y_{est} = Y_p \left[0,0702 - 0,5101 \left(\frac{Y_{aa}}{Y_p} \right) + 4,5968 \left(\frac{ETr}{ET_0} \right) - 0,5514 \left(\frac{ETr}{ET_0} \right) - 0,5471 \left(\frac{ETr}{ET_0} \right) \right] \quad (03)$$

em que:

Y_{est} = Produtividade estimada; Y_p = Produtividade potencial tomada como a máxima dentro da série; Y_{aa} = Produtividade do ano anterior (estimada a partir da equação 03); ETr = Evapotranspiração real; ET₀ = Evapotranspiração potencial; ETr/ET₀ = Evapotranspiração relativa.

3.3. Estimativa da produtividade por lavouras

A avaliação do modelo foi feita a partir de um conjunto de dados de 40 talhões de café entre os anos de 2006 a 2010, obtidos junto à Fazenda Ipanema localizada no município de Alfenas-MG. A produtividade de cada talhão foi comparada à produtividade estimada pelo modelo para os mesmos talhões. Procedeu-se à análise de regressão linear simples com avaliação do modelo pelos intervalos de confiança para o coeficiente de determinação (R²), erro médio absoluto (EMA), erro médio quadrático (EMQ) (Neter et al., 1996) e pela concordância expressa pelo índice d₁ (Willmott et al.; 1981) e índice d₂ (Willmott et al., 1985). O índice d, com variação entre 0 e 1, indica o grau de exatidão entre valores estimados e observados, enquanto o coeficiente de determinação (R²) indica a precisão do modelo, ou seja, quanto da variação da variável dependente é explicada por aquela das variáveis independentes.

A Tabela 02 apresenta os parâmetros estatísticos do modelo. O gráfico de dispersão entre produtividade observada e estimada pelo modelo e apresentado na Figura 01.

Tabela 02 - Valores de coeficiente de determinação (R^2), Erro Médio Absoluto (EMA), Erro Médio Quadrático (EMQ), e índices de Willmott d_1 e d_2 , resultantes da regressão linear entre os dados de produtividade observados em campo e os dados estimados pelo modelo.

Parâmetro estatístico	Inf	Sup
	R^2	0,50
EMA	13,73	17,46
EMQ	18,51	23,15
d_1	0,45	0,53
d_2	0,59	0,70

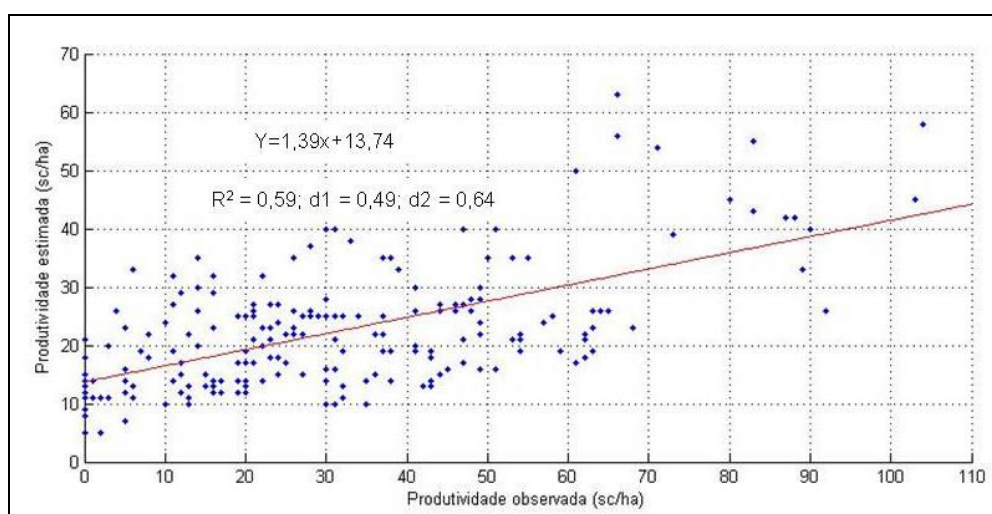


Figura 01 - Diagrama de dispersão entre produtividade observada e estimada utilizando o modelo proposto

De maneira geral, as estimativas em nível de lavouras foram ruins, com coeficientes de determinação (R^2) e índices Willmott (d_1 e d_2) baixos e erros (EMA e EMQ) altos. Em diversos estudos empregando pontualmente modelos semelhantes ao empregado nesta pesquisa, os autores relatam desempenhos superiores (Picinni et al., 1999; Carvalho et al., 2003; Santos et al., 2006). Há de se considerar, entretanto, as limitações inerentes às estimativas de dados meteorológicos espacializados (previsão do ETA) e generalização dos parâmetros para cálculo da evapotranspiração relativa, quando comparados a medições pontuais (estações meteorológicas). Na maioria dos estudos em que foi empregado o modelo agrometeorológico na cultura do café, os dados utilizados foram medidos em estações meteorológicas nas proximidades dos objetos de estudo, ou seja, as lavouras nas quais as produtividades são estimadas.

Rosa (2007) em seu modelo agrometeorológico-espectral utilizou a variável espectral para obter a produtividade potencial da cultura e relata uma diminuição do desempenho das estimativas à medida que o modelo é aplicado em áreas menores com maior detalhamento. Uma provável explicação é exatamente a generalização dos dados meteorológicos como a resolução original dos dados ETA de 40 km, além das limitações inerentes ao modelo, que gera essencialmente uma estimativa. Sugawara (2010), ao avaliar o potencial dos dados ETA em comparação com dados de estações meteorológicas, observou que o modelo superestima os valores de precipitação e apresenta também diferenças significativas entre os valores de temperatura máxima e mínima, comparado ao obtidos nas estações meteorológicas, sendo

estas subestimadas. Tal fato se deve à grande variabilidade espacial das variáveis meteorológicas dentro de uma célula de 40 x 40 km. Entretanto, as possibilidades de evolução na modelagem de dados climáticos resultando em melhores estimativas podem contribuir para modelagem de estimativas da produtividade espacialmente distribuída.

Outra explicação para os resultados ruins é a fragilidade dos modelos empíricos derivados dos índices de vegetação. Os modelos empíricos empregados a partir de índices de vegetação derivados dos dados TM/Landsat para incorporação da variável espectral (estimativa da produtividade do ano anterior) também foram limitados (Bernardes et al., 2013) e (Bernardes, 2013).

3.4. Estimativa da produção total por município

Para esta análise, à semelhança da análise realizada ao nível de lavouras, foram obtidos os parâmetros: coeficiente de determinação (R^2), erro médio absoluto (EMA), erro médio quadrático (EMQ) (Neter et al., 1996) e pela concordância expressa pelo índice d_1 de Willmott et al. (1981) e índice d_2 (Willmott et al., 1985).

Foram somadas as produções de 50 municípios do Sul/Sudoeste de Minas Gerais que estavam totalmente contidos na área de abrangência da imagem Landsat (cena 219/75) entre os anos 2006 a 2010, totalizando 250 observações.

A Tabela 03 apresenta os parâmetros estatísticos do modelo. O gráfico de dispersão entre produtividade observada e estimada pelo modelo e apresentado na Figura 02.

Tabela 03 - Valores de coeficiente de determinação (R^2), Erro Médio Absoluto (EMA), Erro Médio Quadrático (EMQ), e índices de Willmott d_1 e d_2 , resultantes da regressão linear entre os dados de produção total por município e os dados estimados pelo modelo.

Parâmetro estatístico	Inf	Sup
	R^2	0,86
EMA	1607	2147
EMQ	2490	3526
d_1	0,78	0,83
d_2	0,92	0,96

Embora as estimativas ao nível de lavouras tenham apresentado baixo desempenho, ao nível de municípios o modelo se comportou melhor. Os erros variaram entre 1.607 a 2.147 kg (EMA) e 2.490 a 3.526 kg (EMQ). Os valores de R^2 e índices d também foram maiores, variando entre 0,86 a 0,93 para R^2 , 0,78 a 0,83 para o índice d_1 e 0,92 a 0,96 para o índice d_2 . Vale ressaltar que os resultados da Tabela 03 referem-se à produção total por município e não produtividade.

A conversão da produtividade, parâmetro estimado pelo modelo, em produção total por município contribuiu para as melhores correlações com os dados do IBGE, uma vez que a produção total incorpora a área nas estimativas (produtividade X área). Tomando-se o mapa de áreas cafeeiras do ano de 2007 como referência e admitindo-se a sua exatidão, a área de café por município passa a ser um fator importante na produção total, incorporando a exatidão do mapa de áreas cafeeiras à qualidade das estimativas de produção total.

Isto sugere que mesmo utilizando-se o mapa de áreas cafeeiras relativo ao ano de 2007 para estimativas de produção total por município nos anos de 2006 e 2008 a 2010, as correlações permaneceram altas. Tal fato se deve à característica perene da cultura do café e a pequena alteração no parque cafeeiro entre os anos. Moreira (2011) mostra que esta pequena alteração na área de café entre os anos permite um período de monitoramento de até 3 anos

como satisfatório para mapeamento da cultura, ainda assim neste período a variação na área em café não passa de 6%.

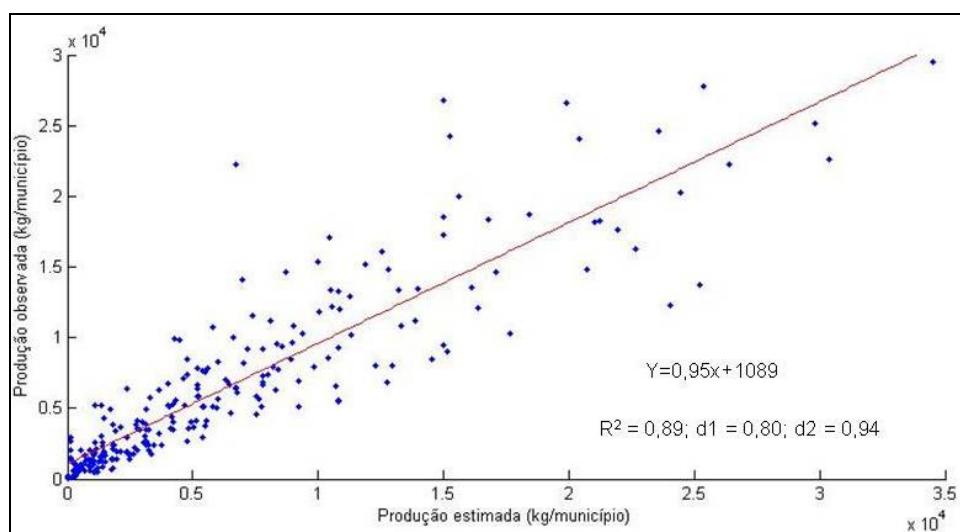


Figura 02 - Diagrama de dispersão entre produtividade observada e estimada

4. Conclusões

No âmbito de lavouras, a estimativa da produtividade por meio da integração de dados TM/Landsat ao modelo agrometeorológico apresentou coeficiente de determinação e índices Willmott baixos e erros altos, o que indica que o modelo não se mostrou satisfatório.

Ao avaliar a produção total por municípios, o modelo apresentou resultado melhor, evidenciando a contribuição positiva do mapa de área plantada na conversão da produtividade em produção.

Referências Bibliográficas

Bernardes, T.; Moreira, M. A.; Adami, M.; Giarolla, A.; Rudorff, B. F. T. Monitoring Biennial Bearing Effect on Coffee Yield Using MODIS Remote Sensing Imagery. **Remote Sensing**, v. 4, n. 9, p. 2492-2509, 2012.

Bernardes, T. **Modelagem de dados espectrais e meteorológicos para estimativa da produtividade de café**. 2013. 128 p. (sid.inpe.br/mtc-m19/2013/03.04.20.07-TDI). Tese (Doutorado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2013.

Bernardes, T., Moreira, M. A., Verona, J. D., Shimabukuro, Y. E., Luiz, A. J. B. Variáveis e modelos para estimativa da produtividade do cafeeiro a partir de índices de vegetação derivados de imagens Landsat. In: Simpósio Brasileiro de Sensoriamento Remoto (SBSR), 16, 2013, Foz do Iguaçu. **Anais...** São José dos Campos: INPE, 2013. Artigos, p. 0720-0727. CD-ROM, On-line. ISBN 978-85-17- 00066-9. Disponível em: <<http://www.dsr.inpe.br/sbsr2013/files/p1437.pdf>>. Acesso em: 10 nov. 2014.

Brunsell, N. A., Pontes, P. P. B. e Lamparelli, R. A. C. Remotely sensed phenology of coffee and its relationship to yield. **GIScience and Remote Sensing**, v.46, n.3, p. 289-304, 2009.

Camargo, M. B. P.; Santos, M. A.; Pedro Junior, M. J.; Fahl, J. I.; Brunini, O.; Meireles, E. J. L.; Bardin, L. Modelo agrometeorológico de monitoramento e de estimativa de quebra de produtividade como subsídio à previsão de safra de café (*Coffea arabica* L.): resultados preliminares. In: SIMPÓSIO DE PESQUISA DOS CAFÉS DO BRASIL, 3., 2003. Porto Seguro. **Anais...** Porto Seguro: Consórcio Brasileiro de Pesquisa e Desenvolvimento do Café, p. 75-76, 2003.

Carvalho, L. G.; Sedyama, G. C.; Cecon, P. R. e Alves, H. M. R. Avaliação de um modelo agrometeorológico para a previsão de produtividade de café em três localidades da Região Sul do Estado de Minas Gerais. **Revista Brasileira de Agrometeorologia**, Santa Maria, v. 11, n. 2, p. 343-352, 2003.

Doorenbos, J. e Kassam, A. H. **Yield response to water**. Rome, FAO, 1979. 197p. (FAO Irrigation and Drainage Paper, 33)

Funk, C. e Budde, M. E. Phenologically-tuned MODIS NDVI-based production anomaly estimates for Zimbabwe. **Remote Sensing of Environment**, v.113, n. 1, p.115-125, 2009.

Instituto Brasileiro de Geografia e Estatística (IBGE). Levantamento sistemático da produção agrícola. 2003. Disponível em: <www.ibge.gov.br/home/estatistica/indicadores/agropecuaria/lspa/>. Acesso em: 31 julho. 2009.

Moreira, M. A.; Barros, M.A.; Faria, V. G. C.; Adami, M. Tecnologia de informação: imagens de satélite para o mapeamento de áreas de café de Minas Gerais. **Informe Agropecuário** (Belo Horizonte), v. 28, p. 27-37, 2007.

Moreira, M. A., Adami, M., Bernardes, T., Rudorff, B. F. T. Análise da dinâmica espaço-temporal da área cafeeira na região sul/sudoeste de Minas, **Anais... XV Simpósio Brasileiro de Sensoriamento Remoto - SBSR**, Curitiba, PR, Brasil, 30 de abril a 05 de maio de 2011, INPE p.0489.

Neter, J.; Kutner, M.H.; Nachtsheim, C.J. e Wasserman, W. **Applied linear statistical models** New York: McGraw-Hill, 1996. 1408p.

Picini, A.G.; Camargo, M.B.P.; Ortolani, A.A. e Gallo, P.B. Desenvolvimento e teste de modelos agrometeorológicos para a estimativa de produtividade do cafeeiro. **Bragantia**, v. 58, n. 1, p. 157-170, 1999.

Rosa, V. G. C. **Modelo agrometeorológico-espectral para monitoramento e estimativa da produtividade do café na região Sul/Sudoeste do estado de Minas Gerais**. (INPE-14791-TDI/1234) Dissertação (Mestrado em Agronomia). Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2007, 143p.

Santos, M. A. e Camargo, M. B. P. Parametrização de modelo agrometeorológico de estimativa de produtividade do cafeeiro nas condições do estado de São Paulo. **Bragantia**, v. 65, n.1, p.173-183, 2006.

Silva, F. M.; Alves, M. C.; Souza J. C. S.; Oliveira, M. S. Efeitos da colheita manual na bienalidade do cafeeiro em Ijaci, Minas Gerais. **Ciência e Agrotecnologia**, v. 34, n. 5, p. 625-632, 2010.

Sugawara, L. M. **Variação interanual da produtividade agrícola da cana-de-açúcar por meio de um modelo agrônomico**. 2010. 114 p. (sid.inpe.br/mtc-19/2010/10.08.14.39-TDI). Tese (Doutorado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2010.

Willmott, C. J. On the validation of models. **Physical Geography**, v.2, p.184-194. 1981.

Willmott, C.J.; Davis, R.E.; Feddema, J.J.; Klink, K.M.; Legates, D.R.; Rowe, C.M.; Ackleson, S.G. e O'donnell, J. Statistics for the evaluation and comparison of models. **Journal of Geophysical Research**, v.90, n.C12, p. 8995-9005, 1985.