

MINERAÇÃO DE DADOS ESPECTRAIS PARA MODELAGEM DE OCORRÊNCIA DE CERCOSPORIOSE EM CAFEZEIROS¹

Margarete M. L. Volpato²; Helena Maria R. Alves³; Tatiana G. C. Vieira⁴; Lívia Naiara de Andrade⁵; Wilian L. Soares⁶; Vanessa Cristina O. Souza⁷; Miguel Thiago Alvarenga⁸; Miler G. Boell⁹

¹ Trabalho financiado pelo Consórcio Brasileiro de Pesquisa e Desenvolvimento do Café – Consórcio Pesquisa Café

² Pesquisadora, Doutora, EPAMIG/URESM, Lavras-MG, BIPDT/ FAPEMIG, margarete@epamig.ufla.br

³ Pesquisadora, PhD, Embrapa Café, Brasília-DF, helenam@embrapa.br

⁴ Pesquisadora, MSc., EPAMIG/URESM, Lavras-MG, BIPDT/ FAPEMIG, tatiana@epamig.ufla.br

⁵ Professora, UFLA/DCC, Lavras-MG, livia.naiara.andrade@gmail.com

⁶ Professor, UFLA/DCC, Lavras-MG, lacerda@dcc.ufla.br

⁷ Professora, UNIFEI/ ICE/DMC, Itajubá-MG, vanessa.vcos@gmail.com

⁸ Bolsista Consórcio Pesquisa Café, BIC, miguelthiago.ufla@gmail.com

⁹ Bolsista Consórcio Pesquisa Café, DCI-2, milergrudtner@gmail.com

RESUMO: O monitoramento fitossanitário possibilita prever o aparecimento ou aumento de intensidade da cercosporiose (*Cercospora coffeicola* Berkeley & Cooke) em cafeeiros. Tradicionalmente esse monitoramento é baseado em observação de períodos críticos ocorridos. Entretanto uma das maiores dificuldades para se utilizar esse tipo de monitoramento é a aquisição de dados climáticos. Uma alternativa para superar este problema é utilizar dados e produtos de imagens de satélites, em função da cobertura espacial e temporal, e de sua relação com as variações do clima e da vegetação de uma região. Uma das dificuldades para realização desse estudo é o grande número de dados gerados, por isso optou-se pela metodologia de mineração de dados, etapa principal do processo de Descoberta de Conhecimento em Bancos de Dados (*Knowledge Discovery in Databases* - KDD). O presente estudo objetivou aplicar técnicas de mineração de dados para encontrar modelos de dados climáticos e espectrais associados à ocorrência da Cercosporiose em cafeeiros. As coletas de dados de campo foram realizadas na fazenda experimental da EPAMIG, em e São Sebastião de Paraíso, MG, e os dados espectrais foram adquiridos pelo sensor MODIS do satélite Terra. Os modelos gerados mostraram que a temperatura média foi o atributo de maior separabilidade na totalidade dos dados climático estudados com taxa de acerto de 67%.

PALAVRAS-CHAVE: Sensoriamento remoto, *Cercospora coffeicola*, café, agrometeorologia.

ABSTRACT: The phytosanitary monitoring enables to predict the onset or increase of intensity of cercospora (*Cercospora coffeicola* Berkeley & Cooke) in coffee. Traditionally this monitoring is based on the observation of critical periods of occurred. However one of the major difficulties in using this type of monitoring is to acquire weather data. An alternative to overcome this problem is to use data and satellite imagery products, depending on the spatial and temporal coverage, and its relationship to changes in climate and vegetation of an area. One of the difficulties in carrying out this study is the large number of generated data, so we opted for the methodology of data mining, the main stage of the process of Knowledge Discovery in Databases (KDD). This study aims to apply data mining techniques to find models of climate data and spectral associated with the occurrence of *Cercospora* leaf spot on coffee. The collected field data were conducted at the experimental farm of EPAMIG in São Sebastião do Paraíso, MG, and the spectral data were acquired by the MODIS sensor Terra satellite. The models showed that the average temperature was higher separability with hit rate of 67%.

KEY WORDS: Remote sensing, *Cercospora coffeicola*, coffee, agrometeorology.

INTRODUÇÃO

O monitoramento fitossanitário da cafeicultura possibilita prever o aparecimento ou aumento na intensidade de doenças, baseado em informações sobre o ambiente, a cultura e/ou o patógeno (Chalfoun, 1997). Tradicionalmente é baseado em observação de períodos críticos das doenças associados às condições meteorológicas, entretanto a maior dificuldade para implementação desse monitoramento está relacionado à aquisição de dados meteorológicos locais. As estações de superfície são mal distribuídas e nem sempre os dados são disponibilizados. Deve-se considerar, ainda, que a grande variabilidade espacial e temporal destes dados impede que os dados locais representem adequadamente uma região. Uma alternativa para superar este problema é utilizar dados e produtos de imagens obtidas por satélites, em função da excelente cobertura espacial e temporal, e de sua relação com as variações do tempo e da vegetação de uma região. Os processos de interação da radiação eletromagnética com a vegetação podem ser estudados segundo as mais diferentes abordagens, incluindo aquelas que consideram a influência de parâmetros específicos das composições química e morfológica de folhas, e ainda daqueles ligados aos demais parâmetros biofísicos da vegetação (Barrett e Curtis, 1992). Neste contexto, estão incluídos os estudos ligados à fitossanidade, nos quais as alterações nos processos fisiológicos

podem exercer influência na interação da radiação eletromagnética com a vegetação. A magnitude desta influência deve ser levada em consideração na definição tanto do nível de aquisição dos dados, quanto das características do sensor a ser utilizado. Para tanto, antes de qualquer iniciativa no sentido de valer-se de técnicas de sensoriamento remoto para a identificação e/ou avaliação de problemas ligados à fitossanidade, é fundamental o conhecimento sobre a influência dos agentes estressantes nos processos fisiológicos e as formas de manifestação dos seus efeitos sobre a interação da radiação eletromagnética com a vegetação (Ponzoni, 1996). O programa Earth Science Enterprise lançado pela NASA (National Aeronautics and Space Administration) em 1991 teve como desafio o desenvolvimento de satélites e outras ferramentas para o estudo intensivo do planeta. Um dos objetivos do programa era oferecer suporte para o monitoramento da agricultura e das florestas, disponibilizando imagens de sensores orbitais capazes de permitir inferências sobre a dinâmica espaço-temporal da superfície da Terra. Assim, no ano de 1999, ocorreu o lançamento do primeiro sensor MODIS (MODerate resolution Imaging Spectroradiometer) que passou a disponibilizar o um novo Índice de Vegetação (IV), denominado Enhanced Vegetation Index (EVI 2). A disponibilização rápida e ininterrupta das imagens MODIS permite análises regulares da vegetação em nível global, incluindo a avaliação de aspectos fenológicos da vegetação e a detecção de mudanças de uso e cobertura do solo. A estreita correlação entre a dinâmica espaço-temporal da vegetação e os IVs, confirma o potencial dos mesmos no monitoramento da cobertura vegetal. No Brasil são escassos os exemplos de aplicação do sensoriamento remoto na identificação e no monitoramento de pragas e doenças, restringindo-se às situações nas quais os plantios já teriam sido parcial ou completamente destruídos. Diante deste problema o presente estudo objetiva aplicar técnicas de mineração de dados para encontrar modelos de dados climáticos e espectrais associados à infestação da doença cercosporiose em cafeeiros. Segundo Pozza e Alves (2008) a cercosporiose é uma doença de grande importância, responsável por perdas significativas na produção de café, e vários estudos já foram realizados a fim de entender a influência do ambiente e do cafeeiro, entretanto existe ainda grande demanda para modelar, prever a ocorrência e formular estratégias de controle dessa doença.

MATERIAL E MÉTODOS

Para este estudo utilizou o banco de dados da Epamig referente ao monitoramento fitossanitário descrito por Souza et al. (2013) conduzido na Fazenda Experimental de São Sebastião do Paraíso, MG, constituído por plantas de cafeeiro arábica, nos anos de 2002 a 2007. Para padronizar e quantificar a infestação da doença cercosporiose em cafeeiros Souza et al. (2013) propuseram a classificação exposta na Tabela 1. Os estágios fenológicos utilizados foram os sugeridos por Camargo & Camargo (2001) (Figura 1). Os dados climáticos foram coletados na estação meteorológica convencional presente na fazenda experimental da EPAMIG, em parceria com a INMET e foram analisados os seguintes dados: precipitação, temperatura mínima, temperatura máxima, umidade relativa do ar. A discretização dos dados de temperatura e precipitação diária seguiu sugestão de Pozza e Alves (2008). Baseados em consultas, foram realizadas entrevistas com os especialistas da área, que sugeriram as seguintes classes: Precipitação Mensal: < 70 e ≥ 70 mm. Umidade: $< 60\%$ e $\geq 60\%$ (Souza et al., 2013). Os dados espectrais foram adquiridos do sensor MODIS do satélite Terra. Os dados espectrais selecionados foram os dados de índice de vegetação EVI2 - Enhanced Vegetation Index 2 disponibilizados por INPE (2011). Foi realizada a avaliação e organização dos dados de índices de vegetação (EVI2) para inserção no Banco de Dados Geográfico (BDG) e em seguida a limpeza dos dados. Valores identificados como erros foram removidos da base de dados. Os dados de EVI2 para o período são apresentados na Figura 2. Os dados de infestação, climáticos e espectrais foram organizados e padronizados quanto a tamanho, formato e domínio, em seguida foi iniciada a etapa de pré-processamento dos dados espectrais EVI2, com a discretização da série temporal de dados EVI2. Esta discretização foi realizada segundo as seguintes classes: $EVI > 0,5$ (dossel vigoroso); $EVI 0,5 - 0,4$ (dossel com vigor vegetativo intermediário); $EVI 0,4 - 0,3$ (dossel com vigor vegetativo baixo); $EVI < 0,3$ (dossel com vigor vegetativo muito baixo). Após a discretização os dados foram encaminhados para a fase de mineração de dados. Os dados foram processados utilizando o software Weka (versão 3.6.1), que é um *software open source* de uma coleção de algoritmos de mineração de dados. Foi utilizado sistema operacional Windows 7. O algoritmo utilizado foi o J48 (Witten & Frank, 2005), que é um indutor de árvore de decisão. Fez se uso do valor 0,3 para intervalo de confiança e *crossover* de 15.

Tabela 1. Classificação da infestação de cercosporiose, em área de monitoramento fitossanitário, conduzido na Fazenda Experimental de São Sebastião do Paraíso, MG.

Nível de Infestação	Número de folhas infestadas
Baixo	1 a 7
Médio	8 a 15
Alto	maior que 15



Fig. 1. Fenologia do Cafeeiro. Fonte: Camargo & Camargo (2001).

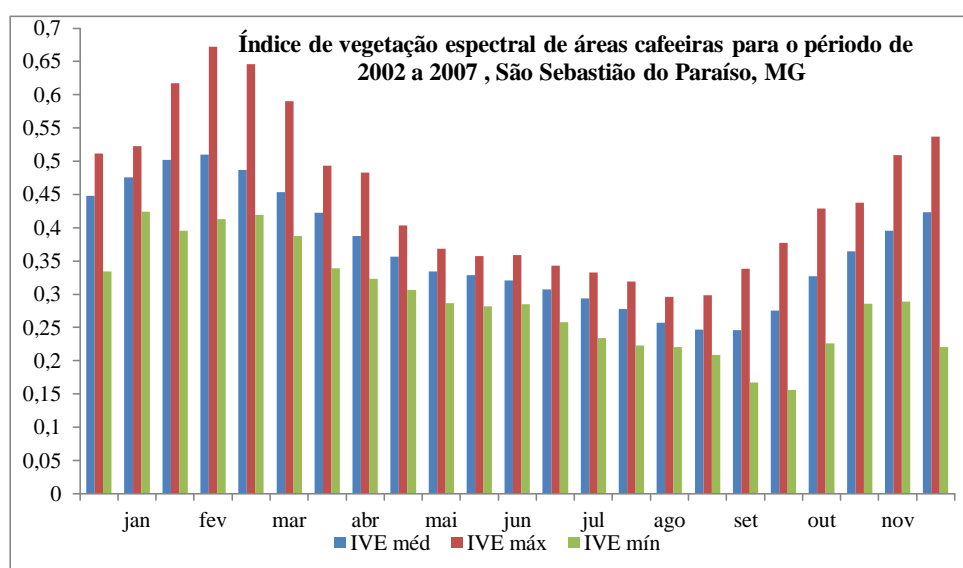


Fig. 2. Série temporal de dados espectrais (EVI 2 - Enhanced Vegetation Index 2).

RESULTADOS E DISCUSSÃO

A árvore de decisão gerada (Figura 3) resultou em uma taxa de acerto de 67%. A leitura da árvore faz-se a partir do topo da árvore (nó raiz) em direção aos nós folhas, passando pelos seus ramos, de acordo com os testes nos valores dos atributos. Os nós folhas representam o objeto de estudo, ou seja, a taxa de infestação da cercosporiose em folha de cafeeiros. Cada ramo na árvore é uma conjunção de condições. Assim, o percurso na árvore (da raiz à folha) corresponde a uma regra de classificação. Verificou-se que a temperatura média foi o atributo de maior separabilidade. No ramo da árvore em que a temperatura média mensal aparece entre 12 e 18 e entre 24 e 30 graus célsius o atributo de separabilidade foi a precipitação média mensal e ocorreu alta infestação quando essa precipitação foi menor que 3 mm e a cultivar avaliada era a Rubi. Pozza e Alves (2008) classificaram como média a favorabilidade da cercosporiose em condições em que as temperaturas média mensal estejam entre 12 e 18 e entre 24 e 30 graus célsius e a precipitação média mensal se mantenha maior que 2 e menor ou igual a 3mm. No ramo da árvore em que a temperatura média mensal aparece entre 18 e 24 graus célsius, a fenologia foi um atributo decisivo para a ocorrência da doença. Pozza e Alves (2008) classificaram como alta a favorabilidade da cercosporiose em condições em que as temperaturas média mensal estejam entre 18 e 24 graus célsius e a precipitação média mensal se mantenha maior que 3 mm. Na fase de Florada, a infestação foi baixa, nas demais fases fenológicas houve pelo menos um conjunto de atributos que resultou em alta infestação da cercosporiose. O atributo EVI2 foi importante para separabilidade nas fases de Repouso e Senescência e Maturação. Na fase fenológica de Maturação o atributo EVI2, na classe 0,3 a 0,4 e no manejo convencional, a infestação foi alta e para o manejo orgânico e cultivar híbrido a infestação também foi alta. Na fase fenológica de Repouso e Senescência o atributo EVI2, na classe 0,4 a 0,5 e no manejo convencional a infestação foi alta. Comparando-se a árvore gerada por Souza et al. (2013) pode-se afirmar que a inclusão do atributo espectral EVI2 não alterou a estrutura da árvore houve porém acréscimo na complexidade da árvore e aumento de 7% na acurácia. De acordo com Souza et al. (2013) o aprimoramento na coleta de dados para construção e validação de novos modelos, pode proporcionar um melhor entendimento da doença por parte dos pesquisadores, técnicos e produtores.

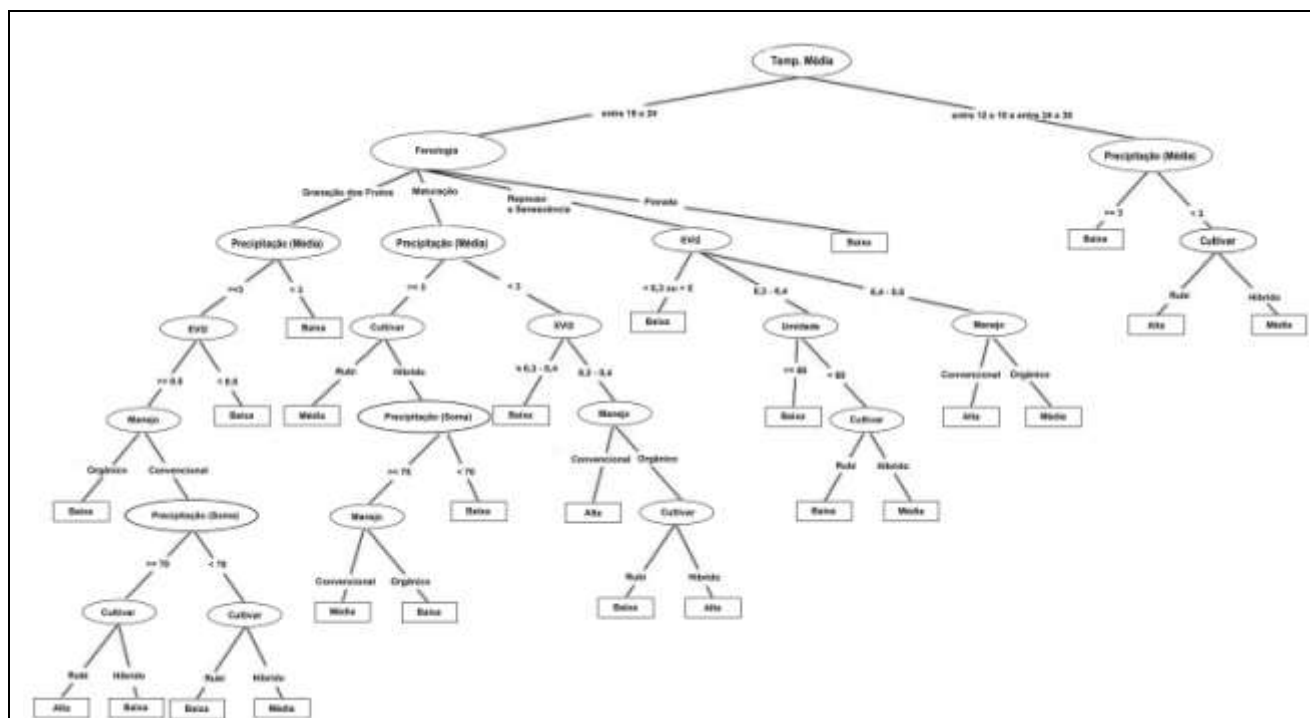


Fig 3. Modelo para ocorrência de cercosporiose em São Sebastião de Paraíso, MG.

CONCLUSÕES

O modelo gerado para explicar a infestação de cercosporiose em folhas de cafeeiros no sul de Minas Gerais, utilizando dados espectrais do satélite Modis, dados climáticos e fenológicos apresentou taxa de acerto de 67%. O modelo gerado possui maior precisão e complexidade quando comparado ao modelo proposto por Souza et al. (2013). A classificação baseada em indução de árvores de decisão no desenvolvimento de modelos que elucidem as variáveis que influenciam na infestação de doenças em cafeeiros é uma técnica nova que necessita de aprimoramento.

REFERÊNCIAS BIBLIOGRÁFICAS

- BARRETT, E. C.; CURTIS, L. F. **Introduction to Environmental Remote Sensing**. 3.ed. London, Chapman & Hall. 1992. 412p.
- CAMARGO, A. P.; CAMARGO, M. B. P. Definição e esquematização das fases fenológicas do cafeeiro arábica nas condições tropicais do Brasil. **Bragantia**, Campinas, v. 6, n. 1, p. 65-68, jan./abr. 2001.
- CHALFOUN, S. M. **Doenças do cafeeiro**: importância, identificação e métodos de controle. Lavras: UFLA/FAEPE, 1997. 96 p.
- INPE. **Visualização das séries temporais MODIS para análise de mudanças de uso e cobertura da terra**. Disponível em: www.dsr.inpe.br/laf/series. Acesso em: junho de 2011.
- PONZONI, F. J. Dados TM/Landsat na identificação do ataque da “vespa-da-madeira” em plantios de Pinus sp. In: Simpósio Brasileiro de Sensoriamento Remoto, 8., Salvador. **Anais ...** São José dos Campos:INPE, abril 1996, 11p.
- POZZA, E. A.; ALVES, M. C. **Impacto potencial das mudanças climáticas sobre as doenças fúngicas do cafeeiro no Brasil**. In: GHINI, R.; HAMADA, E. (Ed.). **Mudanças climáticas: impactos sobre doenças de plantas no Brasil**. Brasília: EMBRAPA, 2008. p. 216-233.
- SOUZA, V. C. O.; CUNHA, R. L.; ANDRADE, L. N.; VOLPATO, M. M. L.; CARVALHO, V. L.; ESMIN, A. A. A. Técnicas de extração de conhecimentos aplicadas á modelagem de ocorrência da Cercosporiose (*Cercospora coffeicola* Berkeley & Cooke) em cafeeiros na região sul de Minas Gerais. **Coffee Science**, Lavras, v. 8, n. 1, p. 91-100, jan./mar. 2013.