

AVALIAÇÃO DA DISPERSÃO ANEMÓFILA DOS ESPOROS DA QUEIMA DAS FOLHAS DO COQUEIRO COM ÁRVORE DE DECISÃO

ANDREIA VIEIRA SANTOS¹
MAIKO DOS SANTOS CORREIA²
JEFFERSON LUIS DA SILVA COSTA³

RESUMO: Este trabalho avaliou o desempenho da técnica de árvore de decisão na construção de modelos climáticos para explicar a dispersão dos esporos da queima das folhas do coqueiro. Os esporos quantificados dessa doença foram correlacionados com variáveis climáticas de precipitação, temperatura máxima, temperatura mínima, temperatura média, umidade relativa do ar e velocidade do vento, no período de junho de 2002 a maio de 2003, no norte do estado de Sergipe. Observou-se que a umidade relativa do ar, entre 70% e 80%, associada com baixas temperaturas, entre 18° e 23°, influenciou significativamente na dispersão dos esporos disseminadores da doença. A baixa intensidade dos ventos, associada com a umidade elevada do ar também foi favorável à dispersão desses agentes. A precipitação influenciou positivamente na dispersão dos esporos, porém, não foi possível definir o volume favorável. Foi possível constatar que as condições climáticas médias, prolongadas por sete dias, proporcionaram uma melhor descrição do comportamento da dispersão dos esporos da queima das folhas do coqueiro, apresentando correlações em torno de 60%. A técnica de árvore de decisão possibilitou uma análise mais detalhada dos dados comparada com a técnica de regressão linear múltipla. No entanto, a influência exercida pelas variáveis estudadas foi semelhante independente da técnica avaliada.

PALAVRAS-CHAVE: doenças de plantas, mineração de dados, análise multivariada.

EVALUATION OF THE ANEMOPHILOUS DISPERSION OF THE LEAF BLIGHTS' SPORES ON COCONUT WITH DECISION TREE

ABSTRACT: This work evaluated the potential of the technique of decision tree in the construction of climatic models to explain the dispersion of the leaf blight's spores on coconut. The quantified spores of that disease were correlated with climatic variables of the precipitation, maximum temperature, minimum temperature, medium temperature, air's relative humidity and wind's speed on period between June of 2002 and May of 2003, in north of the state of Sergipe. It was observed that the air's relative humidity, between 70% and 80%, associated with low temperatures, between 18° and 23°, significant influenced in the dispersion of the spores circulators of the disease. The low wind's intensity associated with the elevated air' humidity s was also favorable to the those agents' dispersion. The precipitation influenced positively in the spore's dispersion, even so, it was not possible to define the favorable volume. It was possible to verify that the medium conditions climatic, prolonged by seven days, provided the best description of the behavior of the dispersion of the leaf blights' spores on coconut, apresented correlations around 60%. The decision tree technique facilitated a analysis of the data more detailed than the techniques multiple lineal regression. However, the influence exercised by the studied variables was similar independent of the appraised techniques.

KEY-WORDS: diseases of plants, data mining, multivariate analysis.

¹ Analista de Sistemas, Embrapa Tabuleiros Costeiros, andreia@cpatc.embrapa.br

² Biólogo, Companhia de Saneamento de Sergipe - DESO, maikocorreia@yahoo.com.br

³ Fitopatologista, Embrapa Sede, jefferson.costa@embrapa.br

1. INTRODUÇÃO

O coqueiro é uma cultura perene de fundamental importância para a economia nordestina brasileira. Devido à possibilidade de aproveitamento praticamente completo da planta, os subprodutos dessa cultura podem ser utilizados para as mais diversas finalidades, trazendo benefícios significativos para as regiões produtoras. No entanto, a sua produção tem sido limitada devido a uma série de fatores, dentre eles, os fitossanitários.

A queima das folhas é uma das mais importantes doenças do coqueiro causada pelo fungo *Lasiodiplodia theobromae* (Leal et al., 2002). A doença é caracterizada principalmente pelo início da infecção no ápice foliar que após atingir a ráquis progride internamente causando o enegrecimento dos tecidos foliares. Como consequência, as folhas que sustentam os cachos caem antes dos frutos completarem a maturação (Correia et al., 2005).

A queima das folhas tem provocado elevadas perdas na produção comprometendo a geração de emprego e renda em toda a cadeia produtiva do coco, principalmente para os pequenos produtores e comunidades rurais cuja principal fonte de renda advém da exploração dessa cultura. Além disso, a ineficiência no controle da doença também implica em problemas ambientais decorrentes do uso indiscriminado de agroquímicos.

A compreensão dos diversos fatores associados a doenças de plantas é de extrema importância no suporte à decisão para o manejo eficiente dos espaços agrícolas. Em relação à queima das folhas do coqueiro, vários estudos envolvendo fatores climáticos e sistemas de manejo de irrigação já foram realizados tanto para avaliar a severidade da doença (Leal et al., 2002) como para explicar, através de equações de regressão, a incidência da mesma por meio da contagem de esporos capturados por armadilhas (Correia et al., 2005).

Diante do grande volume de informações decorrente da variabilidade espacial e temporal do complexo ecológico, a rapidez e a precisão na obtenção de conhecimentos são imprescindíveis para o desenvolvimento sustentável de um agroecossistema. As novas tecnologias computacionais têm promovido agilidade no tratamento e na análise de dados, possibilitando a recuperação de informações acumuladas nas bases, muitas vezes prejudicadas devido à limitação e/ou ineficiência dos métodos convencionais de análise de dados.

A mineração de dados envolve o uso de ferramentas computacionais de análise visando encontrar padrões ou correlações nos dados, auxiliando os usuários no suporte à tomada de decisão (Quoniam et al., 2001). As técnicas disponíveis nessas ferramentas surgem como alternativas a outros métodos de análise de dados, podendo ser exploradas no estudo de doenças de plantas visando o sucesso no controle das mesmas.

2. OBJETIVO

O objetivo deste trabalho foi avaliar o desempenho da técnica de árvore de decisão na construção de modelos climáticos explicativos para a queima das folhas do coqueiro. Além de verificar a influência exercida pelas variáveis climáticas na dispersão dos esporos disseminadores dessa doença, também pretendeu-se identificar de acordo com os vários agrupamentos analisados, o intervalo mais favorável à incidência desses agentes.

3. MATERIAL E MÉTODOS

Os dados analisados foram obtidos no período de junho de 2002 a maio de 2003, no Distrito de Irrigação do Platô de Neópolis, localizado a 92 km de Aracaju, no norte do estado de Sergipe. Nessa região, o clima é do tipo tropical chuvoso, com verão seco e precipitação pluvial em torno de 1.200mm anuais, com chuvas concentradas nos meses de abril a setembro (Cintra et al., 2004).

A incidência da queima das folhas do coqueiro foi avaliada através da correlação entre os esporos disseminadores da doença, capturados por armadilhas e quantificados em laboratório

por Correia et al. (2005), e variáveis climáticas de precipitação (mm), temperatura máxima (°C), temperatura mínima (°C), temperatura média (°C), umidade relativa do ar (%) e velocidade do vento (km/dia), coletadas na estação meteorológica da ASCONDIR (Associação dos Concessionários do Distrito de Irrigação do Platô de Neópolis).

Os esporos foram capturados por uma armadilha do tipo Burkard, instalada nas entrelinhas de um coqueiral anão a 1,80m de altura do solo (altura média da copa da planta). Nessa armadilha, as amostras de ar coletadas foram depositadas em fitas transparentes, semanalmente retiradas e transportadas para o laboratório de biotecnologia da Embrapa Tabuleiros Costeiros. Essas fitas foram cortadas em sete segmentos, cada um contendo a coleta diária dos esporos capturados e quantificados por meio de um microscópio óptico.

A análise dos dados foi realizada através de vários agrupamentos, construídos com dados coletados em intervalos iguais e cumulativos. Para as variáveis precipitação e quantificação dos esporos foram calculados os somatórios dos intervalos avaliados. Para as demais variáveis foram calculadas as respectivas médias de cada intervalo.

Antes de construir os modelos de previsão com a técnica de árvore, foi realizada uma análise prévia dos dados com o propósito de identificar a natureza da sua distribuição e a necessidade de transformá-los em outras potências. Também foi verificada a existência de correlação simples nos dados, de multicolinearidade entre as variáveis climáticas e de viabilidade estatística mínima para empregar as técnicas de análise multivariada selecionadas.

Construção dos modelos com a técnica de árvore de decisão e regressão linear múltipla

No sentido computacional, uma árvore é uma estrutura gráfica de objetos contendo dados, condições e informações (COELHO, 2005). Essa estrutura (Figura1) foi implementada por ferramentas de mineração de dados visando à construção de modelos preditivos.

A técnica utiliza a estratégia “dividir para conquistar” possibilitando, através do particionamento recursivo do conjunto de treinamento (base de dados), que um problema complexo seja decomposto em subproblemas mais simples a partir da geração de novos subconjuntos de dados, cada vez mais homogêneos em relação à variável resposta.

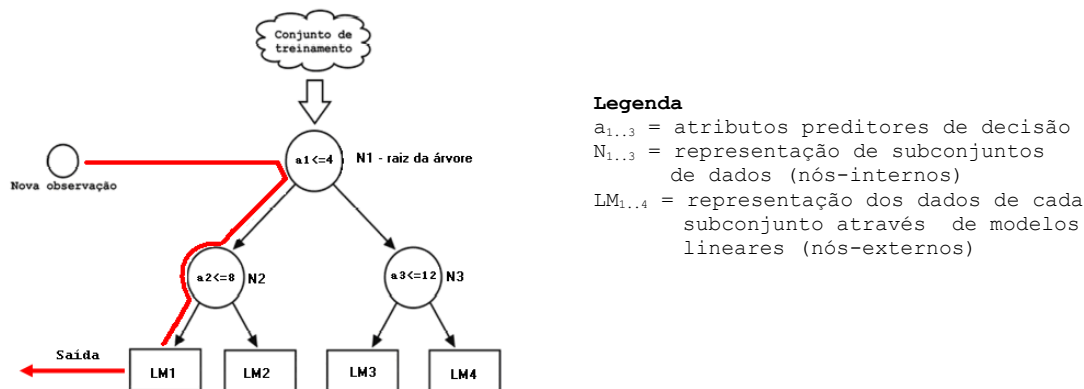


Figura 1. Estrutura de uma árvore de decisão: construção de modelos lineares.

Neste trabalho, as árvores de decisão foram elaboradas através do método *model tree*, implementado pelo algoritmo M5' da ferramenta de mineração de dados WEKA 3.5. Após as sucessivas divisões da base de dados⁴ em novos subconjuntos, os modelos de previsão foram construídos utilizando o método dos mínimos quadrados. Em seguida, esses modelos foram tratados através de procedimentos de suavização e de poda visando reduzir a complexidade

⁴ Visando a máxima redução do erro dos conjuntos de dados analisados, o algoritmo M5' identifica os melhores pontos de divisão desses conjuntos gerando dois novos subconjuntos menos dispersos e mais fáceis de serem modelados: um com valores menores ou iguais ao atributo de decisão e o outro com valores maiores.

das árvores e generalizar os resultados para outras situações próximas das condições climáticas avaliadas.

Para avaliar os modelos foram utilizadas as seguintes medidas de precisão: coeficiente de correlação múltipla (R), erro padrão da estimativa normalizado⁵ (NRMSE) e a técnica de validação cruzada⁶.

Com o objetivo de verificar também a influência exercida pelas variáveis climáticas utilizando métodos tradicionais de análise de dados, foi utilizada a técnica de regressão linear múltipla através dos procedimentos confirmatório e por etapas (*stepwise*).

4. RESULTADOS E DISCUSSÃO

Os dados disponíveis possibilitaram a geração de agrupamentos referentes a intervalos diários e cumulativos até sete dias.

As árvores de decisão foram construídas com todas as possíveis combinações das variáveis climáticas estudadas. No entanto, de acordo com a técnica de validação cruzada, o agrupamento de sete dias cumulativos favoreceu a geração de árvores mais precisas. Os modelos construídos nessas árvores apresentaram correlações moderadas (R entre 0,50 e 0,76) e erros significativos (NRMSE \leq 35%). Na Tabela 1 são apresentados alguns modelos que representaram o conjunto de dados avaliados e os seus respectivos resultados.

Tabela 1. Modelos construídos com o agrupamento de sete dias através da técnica de árvore de decisão.

MODELOS LINEARES	R	NRMSE (%)
$NUM_ESP_PRED_7 = - 515 - 1,8TEMP_MIN_7 + 8,6 UR_7$	0,50	35
$NUM_ESP_PRED_7 = - 401 + 7,2PP_7 + 5,9UR_7$	0,72	29
$NUM_ESP_PRED_7 = - 564 + 9,1UR_7 - 0,3VV_7$	0,60	34

NUM_ESP_PRED = número de esporos preditos; TMIN = temperatura mínima; UR = umidade relativa do ar; PP = precipitação; VV = velocidade do vento; R = coeficiente de correlação múltipla; NRMSE = erro normalizado.

Observou-se que as baixas temperaturas, entre 18° e 23°, apresentaram as melhores correlações com os esporos dispersos da queima das folhas do coqueiro. Segundo Osborne et al. (2001), essas condições também são favoráveis à incidência da giberela do trigo.

A variação da umidade relativa do ar entre 70% e 80% foi favorável à dispersão dos agentes dessa doença, principalmente quando associada às temperaturas apresentadas. A umidade acima de 80% foi citada como influente na dispersão de fungos da giberela do trigo (ROSSI et al., 2002) e da ferrugem do jambeiro (BLUM et al., 2001).

A precipitação, uma das variáveis climáticas já estudadas e identificadas como altamente influente na dispersão dos esporos da queima das folhas do coqueiro (CORREIA et al., 2005), neste trabalho também apresentou influência positiva, porém, não foi possível identificar o volume favorável à dispersão desses microorganismos.

A velocidade do vento exerceu influência negativa na dispersão dos esporos dessa doença, comportamento também verificado quando associada à umidade relativa do ar elevada. Pôde-se concluir que quanto menor a velocidade do vento e maior a umidade relativa do ar, maior foi o número de esporos dispersos.

Em relação aos modelos construídos com a técnica de regressão linear múltipla, não houve diferenças significativas entre os procedimentos confirmatórios e *stepwise*. Da mesma forma que a técnica de árvore, os modelos construídos com o agrupamento de sete dias apresentaram os melhores resultados. A influência exercida pelas variáveis climáticas foi semelhante independente da técnica de análise multivariada empregada. As correlações também foram

⁵ Quanto menor o valor dessa medida, melhor o poder de explicação dos modelos. Um valor acima de 100 indica que a média da variável predita descreve melhor o comportamento dos dados do que os modelos.

⁶ Uma abordagem não paramétrica baseada na reamostragem dos dados e não em suposições estatísticas.

moderadas ($R \cong 60$) com erro médio de previsão em torno de 15% após transformação dos dados com potências logarítmicas.

5. CONCLUSÕES E SUGESTÃO

Constatou-se neste trabalho que as temperaturas baixas, entre 18° e 23°, a umidade relativa do ar, variando entre 70% e 80%, e os ventos com pouca intensidade, denominados de ventos calmos à brisa leve, apresentaram maior influência na dispersão dos esporos da queima das folhas do coqueiro. As condições climáticas prolongadas por sete dias favoreceram a uma maior concentração desses agentes.

Os modelos construídos com a técnica de árvore de decisão possibilitaram uma melhor análise dos dados, mostrando o comportamento dos esporos em diferentes condições climáticas. Outros agrupamentos de dados e técnicas empregando a mineração de dados e métodos tradicionais de análise também poderão ser avaliados no estudo dessa doença.

6. AGRADECIMENTOS

À colaboração dos professores da Universidade Federal de Sergipe; e dos técnicos, analistas e pesquisadores da Embrapa Tabuleiros Costeiros.

7. REFERÊNCIAS BIBLIOGRÁFICAS

- BLUM, L. E. B. B.; DIANESE, J. C. Padrões de liberação de uredinósporos e desenvolvimento da ferrugem do jambeiro. **Pesquisa Agropecuária Brasileira**, Brasília, v. 36, n. 6, p. 845-850, jun. 2001.
- CINTRA, F. L. D.; PORTELA, J. C.; NOGUEIRA, L. C. Caracterização física e hídrica em solos dos Tabuleiros Costeiros no Distrito de Irrigação do Platô de Neópolis. **Revista Brasileira de Engenharia Agrícola e Ambiental**, Campina Grande, v. 8, n. 1, p. 45-50, 2004.
- COELHO, P. S. de S. **Um sistema para indução de modelos de predição baseados em árvores**, 2005. Tese de doutorado. Universidade Federal do Rio de Janeiro. Rio de Janeiro. 252 p.
- CORREIA, M. S.; COSTA, J. L. da S. Dispersão anemófila do fungo *Lasiodiplodia theobromae* em plantações de coqueiro. **Fitopatologia Brasileira**, Brasília, v. 30, n. 2, p. 150-154, mar./abr. 2005.
- LEAL, E. C.; WARWICK, D. R. N.; COSTA, J. L. da S. Doenças do coqueiro: etiologia, amostragem, nível de ação e principais métodos de controle. In: FERREIRA, J. M. S.; MICHEREFF FILHO, M. (eds.). Produção integrada do coco: práticas fitossanitárias. Aracaju, Embrapa Tabuleiros Costeiros. 2002. P. 73-87.
- OSBORNE, L.; JIN, Y. Fusarium head blight: epidemic vs. non-epidemic conditions in South Dakota for 2001. In: NATIONAL FUSARIUM HEAD BLIGHT FORUM PROCEEDINGS. **Anais...** Erlanger, KY, USA, 2001.
- QUONIAM, L.; TARAPANOFF, K.; ARAÚJO JÚNIOR, R. H. de; ALVARES, L. Inteligência obtida pela aplicação de data mining em base de teses francesas sobre o Brasil. **Ciência da Informação**, Brasília, v. 30, n. 2, p. 20-28, maio/ago. 2001.
- ROSSI, V.; LANGUASCO, E.; PATTORI, E.; GIOSUÈ, S. Dynamics of airborne *fusarium* macroconidia in wheat fields naturally affected by head blight. **Journal of Plant Pathology**, v. 84, n. 1, p. 53-64, 2002.