

Mineração de dados para a análise e o alerta da ferrugem asiática da soja

Guilherme Augusto Silva Megeto¹

Carlos Alberto Alves Meira²

Stanley Robson de Medeiros Oliveira²

A ferrugem asiática da soja, causada pelo fungo *Phakopsora pachyrhizi*, é considerada uma das principais doenças que podem limitar a produção da soja, pois, em condições ideais e sem o manejo correto da doença, sua ação é muito severa, o que causa o amarelecimento das folhas e a desfolha precoce, que inviabilizam o enchimento dos grãos de maneira ótima, gerando queda na produtividade e qualidade da colheita, com perdas de mais de 75% da produção (YORINORI et al., 2005). O método de controle da doença mais utilizado e mais eficaz, atualmente, é o uso de fungicidas (GODOY e FLAUSINO, 2008). Porém, o uso de defensivos químicos é oneroso para os produtores, perigoso para a saúde pública e degrada o meio ambiente.

Dentre os métodos utilizados para fazer aplicações mais eficientes de agrotóxicos, destacam-se os modelos epidemiológicos que visam modelar, de maneira qualitativa e quantitativa, as diversas relações presentes no triângulo das doenças – patógeno, hospedeiro e ambiente. Com um melhor conhecimento sobre a doença e seus

¹ Universidade Estadual de Campinas, Faculdade de Engenharia Agrícola; guilhermeasm@cnptia.embrapa.br

² Embrapa Informática Agropecuária; carlos@cnptia.embrapa.br; stanley@cnptia.embrapa.br

fatores mais significativos, o produtor tem o auxílio sobre a aplicação mais eficiente do defensivo químico, se necessária, gerando menor degradação do meio ambiente, menores riscos à saúde e menores gastos com a compra e aplicação dos produtos.

Para estudar o comportamento da ferrugem asiática da soja, alguns modelos já foram propostos e avaliados (DEL PONTE et al., 2006a), mas nenhum desses modelos é absoluto, ou seja, ainda há espaço para melhorias. Os modelos desenvolvidos no Brasil para a avaliação do risco de epidemia e para previsão da severidade final da doença são os modelos desenvolvidos por Del Ponte et al. (2006b), que utilizam regressões, principalmente, da precipitação acumulada em uma determinada janela de tempo.

A proposta deste trabalho é utilizar técnicas de mineração de dados (*data mining*), também conhecida como processo de descoberta de conhecimento em bases de dados (*Knowledge Discovery in Databases, KDD*), para encontrar padrões em bases de dados relacionadas com a cultura da soja, ferrugem asiática e variáveis climáticas, de maneira semelhante à realizada por Meira (2008).

A escolha do processo de mineração de dados para abordar a questão acerca da ferrugem asiática partiu da possibilidade de extrair conhecimentos da base de dados do Consórcio Antiferrugem, que registra todas as ocorrências da doença no Brasil desde 2004 (www.consortioantiferrugem.net), integrada com dados climáticos fornecidos pelo Agritempo (www.agritempo.gov.br). Para realizar o processo de mineração de dados, as ferramentas escolhidas são o software estatístico R para análise exploratória, linguagem de programação Python para manipulação dos dados e os softwares RapidMiner 5.0 e WEKA 3.6 para aplicação de técnicas de mineração de dados.

Até o momento foram realizadas as etapas de compreensão do domínio e entendimento dos dados. Os dados utilizados foram de ocorrências da doença registradas pelo Consórcio Antiferrugem das safras de 2006/07 a 2009/10, com as informações do município, estado, data da ocorrência, tipo de área e estágio fenológico da planta

(VC à R9), totalizando 10.070 ocorrências no Brasil. Dois exemplos de gráficos obtidos na análise exploratória podem ser visualizados na Figura 1.

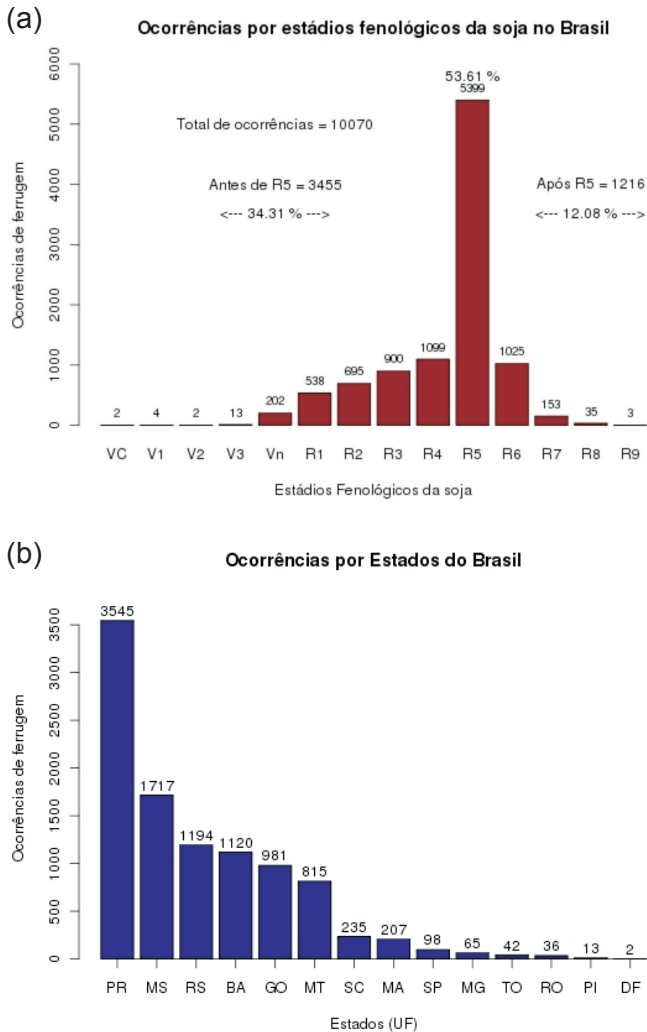


Figura 1. (a) Distribuição das ocorrências por estádios fenológicos. (b) Distribuição das ocorrências por Estados do Brasil.

Alguns aspectos podem ser destacados a partir da visualização dos dados, como o principal estágio de detecção da ferrugem sendo o R5, que corresponde ao enchimento de grãos, e os estados da Bahia, Paraná, Rio Grande do Sul e Mato Grosso do Sul, juntos, correspondem a mais de 75% das ocorrências registradas no país.

Com esses primeiros resultados, duas possibilidades de estudo podem ser derivadas: i) Analisar apenas os dados das regiões mais representativas, no caso, dos quatro estados citados anteriormente; ii) Na modelagem, a possibilidade da utilização de algoritmos que necessitem que os exemplos possuam classes, como os de árvores de decisão, em que o estágio de desenvolvimento da planta pode ser utilizado para derivar o atributo classe, por exemplo, com a criação de três classes: antes de R5, R5 e após R5.

Uma melhor compreensão do domínio e consultas com especialistas devem auxiliar na divisão correta dessas classes, baseando-se nas semelhanças e diferenças de um estágio para outro. Os dados climáticos como temperaturas mínima e máxima e precipitação ainda estão em processo de aquisição.

O resultado esperado é obter modelos que evidenciem os fatores que mais contribuem para a ocorrência da doença, aplicando-se um processo de mineração de dados no conjunto integrado de dados do Consórcio Antiferrugem e meteorológicos.

Referências

- DEL PONTE, E. M.; GODOY, C. V.; CANTERI, M. G.; REIS, E. M.; YANG, X. B.; Models and applications for risk assessment and prediction of Asian soybean rust epidemics. **Fitopatologia Brasileira**, Brasília, DF, v. 31, p. 533-544, Dec. 2006a.
- DEL PONTE, E. M.; GODOY, C. V.; LI, X.; YANG, X. B. Predicting severity of Asian soybean rust epidemics with empirical rainfall models. *Phytopathology*, v. 93, No. 7, p. 797-803, 2006b.

GODOY, C. V.; FLAUSINO, A. M. **Eficiência de fungicidas para o controle da ferrugem asiática da soja em Londrina e Tamarana, PR, na safra 2007/08**. Londrina: Embrapa Soja, 2008. 4 p. (Embrapa Soja, Circular Técnica 57)

MEIRA, C. A. A. **Processo de descoberta de conhecimento em bases de dados para a análise e o alerta de doenças de culturas agrícolas e sua aplicação na ferrugem do cafeeiro**. 2008. 198 p. Tese (Doutorado) – Faculdade de Engenharia Agrícola, Universidade Estadual de Campinas, Campinas. 2008.

YORINORI, J. T.; PAIVA, W. M.; FREDERICK, R. D.; COSTAMILAN, L. M.; BERTAGNOLLI, P. F.; HARTMAN, G. L.; GODOY, C. V.; NUNES-JÚNIOR, J. Epidemics of soybean rust (*Phakopsora pachyrhizi*) in Brazil and Paraguay from 2001 to 2003. **Plant Disease**, v. 89, p. 675-677, June 2005.