

TIC na segurança fitossanitária das cadeias produtivas

Jayme Garcia Arnal Barbedo
Carlos Alberto Alves Meira

1 Introdução

Com a intensificação da indústria agropecuária, têm crescido os desafios e as preocupações relacionadas à segurança sanitária dos alimentos produzidos. A circulação de volumes cada vez maiores desse tipo de mercadoria exige que as medidas necessárias para garantir sua segurança sanitária sejam implementadas de maneira rápida, eficiente e barata. O controle manual tradicionalmente utilizado muitas vezes não é capaz de atender a esses requisitos. Como resultado, tecnologias de informação e comunicação têm sido cada vez mais utilizadas para:

- 1) Aumentar o grau de automação e, conseqüentemente, a velocidade dos processos de controle fitossanitário.
- 2) Identificar problemas sanitários tão cedo quanto possível, minimizando possíveis prejuízos econômicos, ambientais e sociais.
- 3) Identificar, a partir de variáveis ambientais e históricas, áreas potencialmente sujeitas a problemas sanitários, antes mesmo destes se manifestarem.

Este capítulo trata especificamente dos dois últimos itens. Na Seção 2, são mostradas iniciativas voltadas ao diagnóstico de doenças em plantas, explorando tecnologias como processamento digital de imagens e sistemas especialistas. A Seção 3, por sua vez, apresenta iniciativas voltadas à construção de modelos de previsão e sistemas de alerta de doenças de culturas agrícolas.

2 Diagnose de doenças

Desde os primórdios da agropecuária, uma das principais preocupações dos produtores tem sido evitar prejuízos ligados à manifestação e proliferação de doenças em suas lavouras, viveiros e rebanhos. A fim de que o combate a essas doenças seja eficaz, é essencial que o diagnóstico seja não apenas correto, mas tão rápido quanto possível. Na maioria das vezes, o monitoramento das condições sanitárias é feito visualmente, através da ida a campo de especialistas capazes de reconhecer possíveis problemas que estejam ocorrendo. Essa estratégia tem alguns problemas potenciais:

- Especialistas nem sempre estão disponíveis para realizar o monitoramento com a frequência necessária, especialmente em locais remotos.
- O custo associado a esse monitoramento por especialistas pode ser proibitivo para pequenos produtores.
- Ainda que haja disponibilidade de mão de obra e recursos, pode ser muito difícil fazer um monitoramento minucioso, especialmente no caso de grandes propriedades.

Como resultado, esforços vêm sendo despendidos na criação de ferramentas computacionais que auxiliem no combate aos problemas sanitários. O objetivo dessas ferramentas é reduzir o tempo necessário para que as primeiras ações sejam colocadas em andamento, especialmente na ausência de um especialista.

Esta seção apresentará alguns dos principais avanços alcançados nos últimos anos nas técnicas para detecção, quantificação e classificação de doenças na agropecuária, com ênfase naqueles baseados em imagens digitais dos sintomas das doenças, os quais são a grande maioria. Será dado particular destaque às ações que têm sido realizadas na Empresa Brasileira de Pesquisa Agropecuária (Embrapa).

2.1 Sistemas automáticos

Sistemas automáticos são aqueles que realizam todas as operações automaticamente com base nas imagens digitais a eles submetidos, com pouco ou nenhum envolvimento por parte do usuário. Este alto grau de autonomia possui algumas vantagens:

- O usuário não necessita de qualquer conhecimento específico a respeito do problema, já que toda a informação necessária para o diagnóstico está contida no algoritmo.
- Computadores não se cansam, portanto é possível fazer um grande número de avaliações ininterruptamente, o que possibilita monitoramento permanente.
- Computadores não estão sujeitos a ilusões de ótica que frequentemente acometem avaliadores humanos.
- Em geral possuem baixo custo de operação.

A principal desvantagem deste tipo de sistema é sua dependência em relação à qualidade da base de dados de imagens digitais utilizada no seu desenvolvimento. O resultado disso é que, embora seja possível conferir um certo grau de flexibilidade ao algoritmo, este só será capaz de lidar com situações para as quais ele foi treinado. Assim, é possível afirmar que um sistema automático para detecção, medição ou identificação de doenças só poderá ser tão bom quanto a base de dados utilizada. Na verdade, a falta de bases de dados abrangentes é o principal problema enfrentado atualmente nas pesquisas deste tipo, já que as técnicas de processamento de imagens e inteligência computacional estão suficientemente maduras para permitir o desenvolvimento de métodos realmente efetivos. Por esse motivo, há várias iniciativas em andamento para a construção de bases de dados mais completas, inclusive na Embrapa, conforme será descrito mais adiante.

A seguir é apresentado um breve estado da arte, bem como as iniciativas realizadas na Embrapa, relacionados aos temas de detecção, quantificação e classificação de doenças. É importante ressaltar que plantas têm recebido mais atenção nesses temas que animais, motivo pelo qual apenas plantas são tratadas nas seções apresentadas a seguir.

2.1.1 Detecção de doenças

Métodos para detecção de doenças em plantas normalmente são voltados ao monitoramento, em tempo real, da lavoura. Assim, o objetivo é detectar, o mais cedo possível, a presença de uma ou mais doenças capazes de trazer prejuízos à produção.

Embora o monitoramento e detecção de doenças seja de grande importância, não há na literatura muitas propostas voltadas especificamente para este fim. Isso é provavelmente decorrente do fato de que métodos para classificação de doenças, os quais normalmente empregam técnicas mais sofisticadas, se prestem também a este tipo de aplicação.

Dentre os métodos de detecção de doenças em plantas encontrados na literatura, destacam-se aqueles propostos por Sena Junior et al. (2003) e por Story et al. (2010). O primeiro é baseado na segmentação da imagem por limiarização e na contagem dos objetos detectados, enquanto o segundo realiza uma análise de regressão duplamente segmentada com base em atributos de cor e textura. Embora não haja na Embrapa nenhuma iniciativa especificamente dedicada a este fim, dentre os resultados a serem gerados dentro do projeto “Diagnóstico automático de doenças em plantas usando imagens digitais” (BARBEDO, 2013f) está um classificador de doenças que, com algumas poucas adaptações, terá a capacidade de realizar esse tipo de monitoramento. Mais informações a respeito deste projeto, o qual tem financiamento da Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp) e está sob liderança da Embrapa Informática Agropecuária, serão fornecidas mais adiante, na parte de classificação de doenças.

2.1.2 Medição da severidade de doenças

A medição da severidade das doenças é muito importante no contexto do controle sanitário. Em geral, essa medição é feita de maneira subjetiva por especialistas, os quais usam escalas padronizadas para emitir um escore que indica o grau de severidade dos sintomas visíveis. Em alguns casos, faz-se também a medição manual das áreas usando um padrão quadriculado. Pesquisas demonstram que há diferenças significativas nos escores atribuídos pelos especialistas, levando a resultados inconsistentes (BOCK et al., 2010). Além disso, o processo de medição manual das áreas é cansativo, demorado e sujeito a erros. Por esses motivos, a automação da medição da severidade dos sintomas tem recebido grande atenção nos últimos anos. Um levantamento completo dos métodos propostos na literatura pode ser encontrado em Barbedo (2013a), sendo que serão citados aqui apenas aqueles com resultados mais relevantes. Esta seção focará nas três principais maneiras de se medir a severidade das doenças: estimando a área das lesões, contando o número de lesões, ou analisando as características de cor e textura das lesões.

A estimativa da área das lesões é, por uma larga margem, a estratégia mais utilizada para determinar a severidade da doença. Dentre as ferramentas deste tipo recentes, pode-se citar aquelas propostas por Lloret et al. (2011), a qual é baseada em limiarização e morfologia matemática, e por Contreras-Medina et al. (2012), a qual é baseada em análise de cor para realizar a segmentação das lesões. Na Embrapa, a medição da área de lesões foi estudada no contexto do projeto “Contagem semi-automática de objetos genéricos em imagens digitais” (Contag), liderado pela Embrapa Informática Agropecuária. Esses estudos geraram um método para medição de lesões inteiramente baseado em conhecimento especialista e operações morfológicas, método este descrito em detalhes em Barbedo (2013a, 2014b). A Figura 1 apresenta um resumo de todos os passos que compõem o método desenvolvido. Este método será implementado como um módulo do sistema para contagem e medição de áreas de objetos, a ser disponibilizado na internet até o início de 2015.

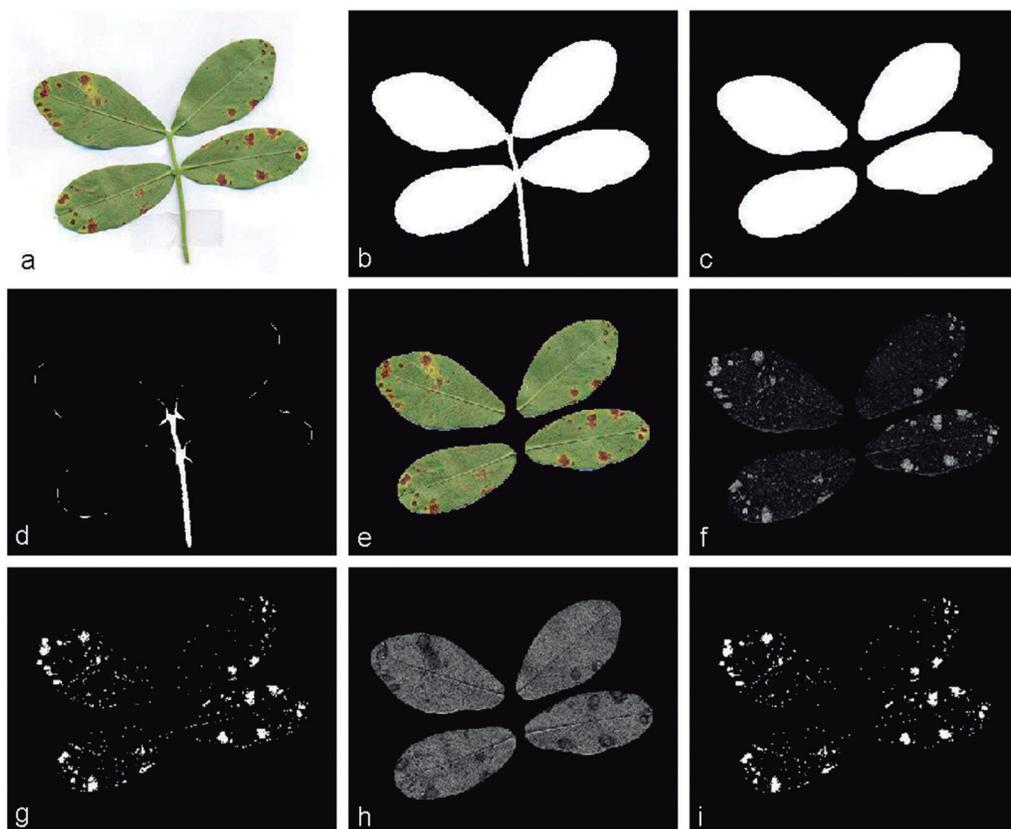


Figura 1. Passos do algoritmo. a) Imagem típica presente na base de dados. b) Máscara binária antes do ajuste fino. c) Imagem aberta morfológicamente. d) Imagem contendo as localizações do caule e do pecíolo antes da eliminação de objetos espúrios. e) Imagem contendo apenas as regiões de interesse. f) Representação da imagem no canal amarelo do espaço de cor CMYK. g) Primeira estimativa para a localização das lesões. h) Representação da imagem no canal magenta do espaço de cor CMYK. i) Estimativa final para a localização das lesões.

A contagem de elementos ou estruturas que caracterizem a doença é também uma abordagem comum. Um exemplo desse tipo de estratégia pode ser encontrado em Boissard et al. (2008). Como mencionado anteriormente, o projeto Contag foi criado exatamente para lidar com o problema da contagem de objetos, embora seu escopo seja mais amplo que a simples contagem de estruturas relacionadas a doenças. Os resultados desse projeto foram aplicados, majoritariamente, a problemas de contagem de microrganismos (BARBEDO, 2012, 2013b, 2013c, 2013d). O algoritmo foi também aplicado à contagem de moscas brancas em folhas, as quais podem causar doenças e são uma importante fonte de perdas na agricultura (BARBEDO, 2014c). A Figura 2 fornece um exemplo de uso deste método, o qual é também inteiramente baseado em regras ad-hoc, geradas a partir de conhecimento de especialistas, e morfologia matemática. Todas essas ferramentas de contagem estarão também disponíveis como aplicativo web a partir do início de 2015.

O uso de características de cor e textura para medir a severidade de doenças é mais incomum que as outras duas estratégias, provavelmente devido às dificuldades em se garantir que as imagens capturem com fidelidade as características visuais dos sintomas, e também devido à falta de ho-

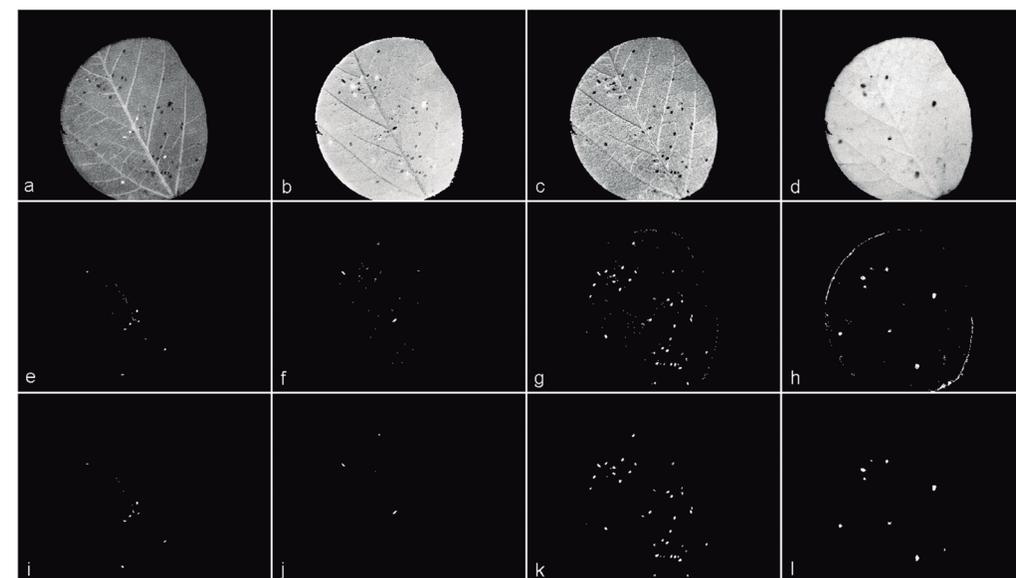


Figura 2. Ilustração dos passos do algoritmo para contagem de moscas brancas em folhas.

mogeneidade dessas características. Até onde vai o conhecimento dos autores, o único método a usar este tipo de abordagem foi proposto por Zhou et al. (2011).

2.1.3 Classificação de doenças

A classificação de doenças é, em geral, um problema mais difícil que aqueles apresentados até aqui, uma vez que, além de detectar a doença, o algoritmo deve tentar identificá-la. O problema se torna mais difícil à medida que um maior número de doenças é considerado. Apesar de haver um grande número de métodos para este fim, a grande maioria deles possui algum tipo de restrição em relação às espécies e doenças consideradas, bem como às condições sob as quais as imagens são capturadas (BARBEDO, 2013a). Dentre os métodos mais bem-sucedidos propostos na literatura, pode-se citar Camargo e Smith (2009), baseado em máquinas de vetores suporte; Huang (2007), baseado em redes neurais do tipo MLP; Pydipati et al. (2006), baseado em análise de discriminante e Xu et al. (2011), baseado em lógica nebulosa.

O projeto Digipathos pretende contribuir para o avanço desta área através de duas ações descritas a seguir.

a) Criação de uma base de dados contendo imagens de doenças de pelo menos 20 espécies com valor comercial no Brasil. Como comentado anteriormente, um método automático de diagnóstico somente pode ser tão abrangente quanto a base de dados usada no seu desenvolvimento, e as pesquisas na área sofrem com a falta de bases de dados estruturadas, completas e com imagens de boa qualidade. A fim de facilitar a submissão das imagens e organizar as imagens enviadas pelos pesquisadores envolvidos no projeto, foi criado um repositório web usando a plataforma DSpace (SMITH et al., 2003). A tela inicial do repositório é mostrada na Figura 3.

Além do repositório, foi criada uma planilha para Excel e OpenOffice, na qual os pesquisadores responsáveis podem, através do clique de um botão, listar todas as imagens presentes no diretório, visualizá-las, e preencher seus metadados.

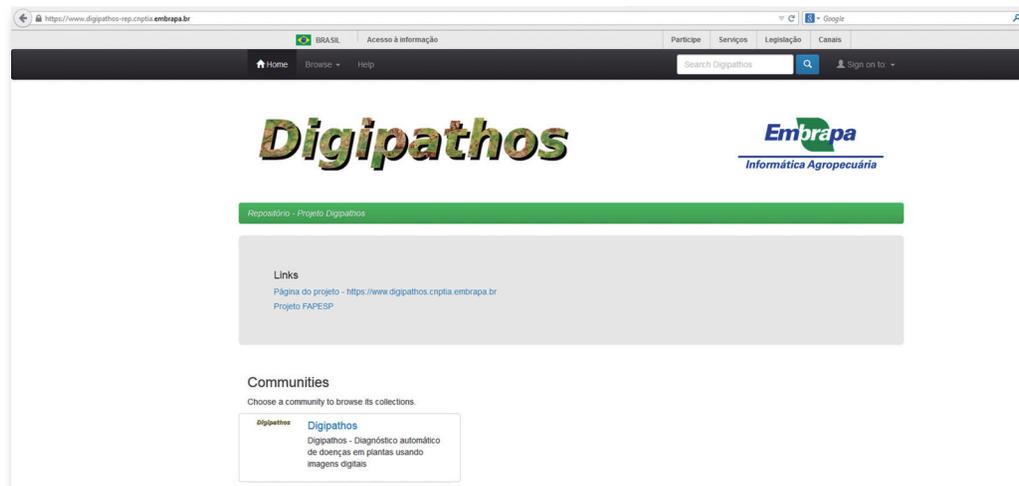


Figura 3. Tela inicial do repositório de imagens de sintomas de doenças em plantas.

Atenção especial foi dada também à captura das imagens. Embora se pretenda utilizar imagens capturadas sob condições normais ao longo do projeto, as etapas iniciais de desenvolvimento do método se beneficiariam de imagens com condições mais controladas de iluminação, e corrigidas em termos de cor e geometria. Por esse motivo, foi criado um dispositivo, consistindo de uma caixa e de um padrão (Figura 4), o qual, além de garantir condições homogêneas de captura, permite que se utilize um software para correções geométricas e de cor das imagens. Detalhes a respeito do dispositivo podem ser encontrados no website do projeto¹.

b) Desenvolvimento de método para identificação de doenças em plantas. Esse método deverá se basear em técnicas de aprendizado de máquina, reconhecimento de padrões, morfologia matemática e conhecimento especialista. O aplicativo baseado nesse método terá duas ver-

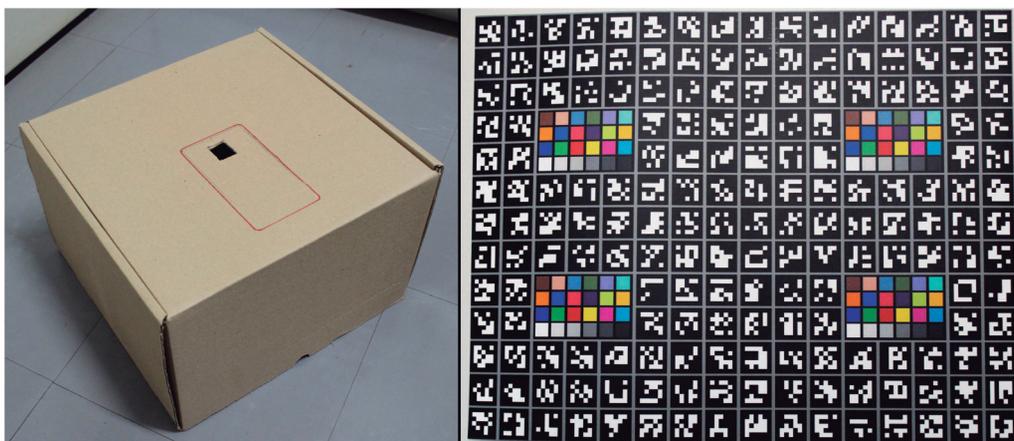


Figura 4. Caixa e padrão usados na captura das imagens de sintomas de doenças, no contexto do projeto Digipathos.

¹ Disponível em: <<https://www.agropediabrasilis.cnptia.embrapa.br/web/digipathos>>.

sões, uma web, a ser usada em computadores com conexão com a internet, e uma móvel, a ser embarcada em aparelhos celulares e tablets, não dependendo da existência de uma rede para funcionar.

2.2 Sistemas semi-automáticos

Sistemas semi-automático são aqueles que, além de terem uma parte computacional que realiza certas operações que levam ao diagnóstico, dependem também de uma participação humana para funcionarem corretamente. Pode-se dividir estes sistemas em dois tipos, os quais são apresentados nas subseções a seguir.

2.2.1 Correção manual dos resultados

Neste tipo de abordagem, a parte automática do sistema fornece uma resposta, a qual, se apresentar falhas ou erros visíveis, pode ser corrigida pelo usuário. Evidentemente, este tipo de abordagem não serve para a classificação de doenças, uma vez que se o usuário sabe que houve erro; então, ele já saberia a resposta de antemão, não havendo necessidade de usar o programa. Contudo, em análises quantitativas é possível, por exemplo, identificar manualmente objetos que o algoritmo não detectou, ou corrigir as bordas das áreas com sintomas detectadas. A interface do programa criado no contexto do projeto Contag permite ao usuário realizar este tipo de correção, conforme exemplificado na Figura 5.

Não é comum, na literatura, encontrar métodos ou sistemas originais que permitam a correção manual por parte do usuário. Por outro lado, ferramentas genéricas de processamento digital de imagens, tais como o ImageJ², o Assess³ e o SigmaScan Pro⁴, naturalmente dependem das

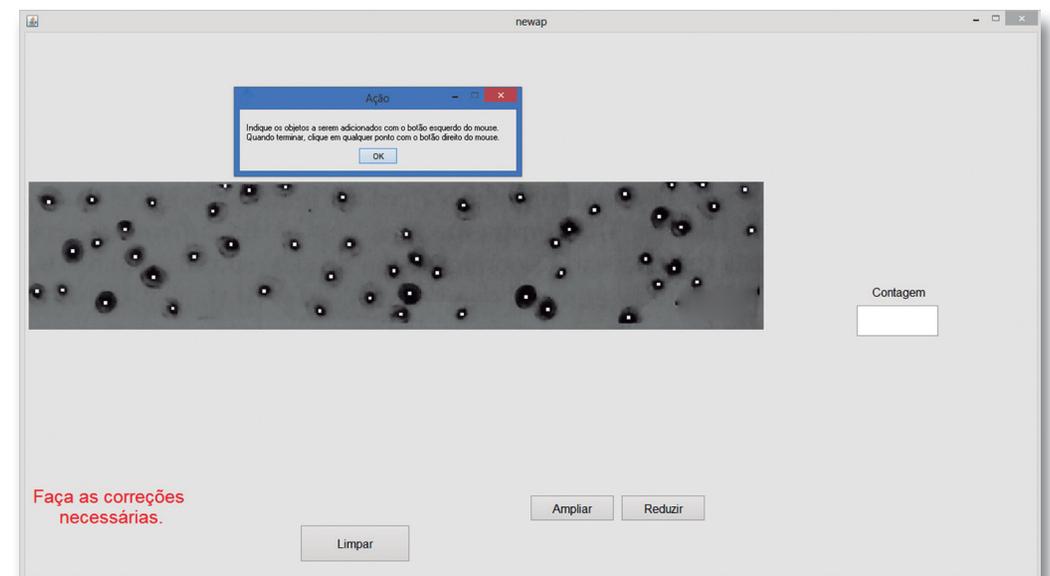


Figura 5. Exemplo de correção manual presente no sistema desenvolvido no contexto do projeto Contag.

² Disponível em: <<http://imagej.nih.gov/ij/>>.

³ Disponível em: <<https://www.apsnet.org/apsstore/shopapspress/Pages/43696m5.aspx>>.

⁴ Disponível em: <<http://www.sigmaplot.com/products/sigmaScan/sigmaScan.php>>.

entradas e intervenções feitas pelos usuários. Tais ferramentas normalmente são aplicadas em uma grande variedade de problemas, incluindo aqueles relacionados à agricultura. Um estudo comparando abordagens manuais, semi-automáticas e automáticas foi apresentado por Barbedo (2013e).

2.2.2 Sistemas especialistas

Sistemas especialistas são ferramentas baseadas em um conjunto de regras que tentam representar, da maneira mais acurada possível, todo o universo de possibilidades do problema que se deseja resolver. No caso específico de doenças em plantas, inicia-se com todo o universo de culturas e doenças para as quais o sistema foi treinado e, através de perguntas, as possibilidades vão sendo sucessivamente refinadas, até que se chegue à resposta desejada.

Este tipo de sistema é bastante popular em diversas áreas de aplicação. Estes começaram a ser utilizados na agricultura na década de 1980. Desde então, um grande número de sistemas foi proposto, como o Plant/ds, para diagnóstico de doenças da soja (MICHALSKI et al., 1982), o MoreCrop⁵ para previsão e manejo da ferrugem, e o TomEx, desenvolvido no Brasil com foco em doenças do tomateiro (POZZA et al., 1997). Na verdade, pode-se encontrar na literatura centenas de sistemas especialistas diferentes, usando as mais diferentes técnicas e abordagens. Normalmente esses sistemas são desenvolvidos tendo uma única cultura particular em mente e, ainda mais importante, o conjunto de regras utilizado normalmente só é válido para as condições climáticas, ambientais e epidemiológicas típicas do país ou região no qual se pretende utilizar o sistema. Como consequência, são raras as comparações entre sistemas especialistas. Esse fato, por sua vez, faz com que não seja possível determinar o estado da arte desta área, já que cada sistema resolve um pequeno subconjunto do problema mais geral do diagnóstico de doenças em plantas.

Na Embrapa, foram desenvolvidos dois sistemas diferentes:

- **Sistema Diagnose Virtual:** desenvolvido sob liderança da Embrapa Informática Agropecuária, este sistema possui uma infraestrutura única na área de sanidade para diagnóstico de doenças de plantas via internet, a fim de subsidiar os agricultores, agrônomos e técnicos agrícolas em suas decisões sobre o manejo de doenças (MASSRUHÁ; LIMA, 2011; MASSRUHÁ et al., 2007a, 2007b, 2007c; SILVA et al., 2011). Visa possibilitar o uso racional de agrotóxicos, o que ajuda a evitar mais danos à saúde e ao meio ambiente, além de reduzir os custos da produção. O sistema possui dois módulos: a) módulo especialista, no qual fitopatologistas têm a responsabilidade de fornecer ao sistema uma base de conhecimento sobre doenças de uma determinada cultura; b) módulo produtor: os produtores rurais fornecem as informações sobre os sintomas encontrados em sua cultura, e obtêm como resultado a indicação das prováveis desordens e as medidas de controle viáveis. Atualmente, o sistema é capaz de fornecer diagnóstico para as culturas do arroz, feijão, milho, soja, tomate e trigo. A Figura 6 mostra a tela inicial do sistema, o qual pode ser acessado pelo endereço⁶
- **Sistema Uzum:** desenvolvido pela Embrapa Uva e Vinho, este sistema visa servir como um guia para auxiliar na identificação de possíveis agentes causadores de sintomas em videiras, fornecendo informação útil para um rápido diagnóstico inicial. Após o usuário fornecer as respostas solicitadas nas telas da interface, o sistema fornece uma lista de possíveis problemas,

⁶ Disponível em: <<http://www.diagnose.cnptia.embrapa.br/diagnose/>>.



Figura 6. Tela inicial do módulo produtor do sistema Diagnose Virtual.

com as respectivas probabilidades. O sistema tem registrados 37 doenças, pragas e distúrbios fisiológicos, estando disponível na internet⁷. Mais detalhes podem ser encontrados em Fialho et al. (2012), e uma imagem da tela inicial é mostrada na Figura 7.



Figura 7. Tela inicial do sistema especialista para diagnóstico de doenças em videiras Uzum.

Normalmente, sistemas especialistas usados para diagnóstico de doenças em plantas precisam considerar uma ampla gama de problemas, fazendo com que seja necessária a criação de um extenso conjunto de regras, as quais devem se relacionar de maneira coerente a fim de resultar em um bom diagnóstico. Um extenso conjunto de regras significa que o processo para se chegar a um diagnóstico acurado pode demandar que um grande número de perguntas seja respondido

⁷ Disponível em: <<http://www.cnpuv.embrapa.br/tecnologias/uzum/>>.

pelo usuário. Este fato aumenta consideravelmente a probabilidade de erros, uma vez que uma única pergunta respondida de maneira errônea pode fazer com que o sistema se desvie irremediavelmente da resposta correta.

Uma maneira de se limitar o universo inicial de possibilidades é fornecer ao sistema especialista informações objetivas capazes de eliminar diagnósticos improváveis antes mesmo do usuário começar a responder as perguntas. Tendo esse objetivo em mente, está em andamento um projeto para acoplar o sistema automático baseado em imagens digitais a ser desenvolvido no projeto Digipathos, com um novo sistema especialista, o qual aproveitará muitos dos conceitos usados no sistema Diagnose Virtual para criar uma ferramenta com abrangência ainda maior. Assim, o sistema baseado em imagens será capaz de reduzir o número de respostas possíveis, e o sistema especialista será responsável por dirimir qualquer incerteza advinda da parte baseada em imagens. O novo sistema deverá ser implementado em duas versões, uma web e uma a ser embarcada em dispositivos móveis.

2.3 Conclusão

O uso da tecnologia para auxiliar no diagnóstico de doenças na agropecuária tem experimentado significativos avanços nas últimas décadas. Dada a importância deste tema, a Embrapa tem adotado diversas iniciativas para desenvolver tecnologias próprias e contribuir para o avanço do tema, como pode ser comprovado por ferramentas como o Diagnose Virtual e o Uzum, e projetos como o Contag e o Digipathos. Contudo, os desafios a serem vencidos ainda são consideráveis, ensejando que novas iniciativas sejam levadas adiante num futuro próximo.

3 Sistemas de alerta de doenças de plantas

Um sistema de previsão de doença de planta é aquele que prevê o aparecimento ou um aumento na intensidade de uma doença baseado em informação sobre o ambiente, a cultura e/ou o patógeno (CAMPBELL; MADDEN, 1990). Esse aparecimento ou aumento futuro da doença é frequentemente baseado na observação de períodos críticos ocorridos, o que acaba causando certa confusão. Com respeito aos sintomas, a previsão é anterior ao fato, mas com respeito à infecção, a previsão é posterior. Zadoks (1984) sugeriu o termo aviso de doença ou alerta de doença, para evitar problemas de terminologia e enfatizar que a mensagem aos produtores é mais importante do que a sua origem técnica.

Prever doenças de plantas é importante por duas razões principais: economia e segurança. A questão econômica é reduzir o custo de produção por meio de aplicações oportunas de medidas de controle, geralmente na forma de fungicidas. Segurança envolve não apenas a cultura, reduzindo efeitos tóxicos sobre as plantas, mas também o ambiente externo, reduzindo a exposição de agrotóxicos a outras espécies de plantas, aos trabalhadores e aos consumidores (HARDWICK, 2006).

Os alertas auxiliam os produtores a determinar a necessidade e o momento de aplicar técnicas de controle de doenças (CAMPBELL; MADDEN, 1990). Um alerta proporciona indicação de quando é provável que a doença vá se tornar crítica e, portanto, ter impacto econômico. Para algumas doenças, é importante ser capaz de prever a primeira ocorrência, enquanto para outras um certo nível de doença pode ser tolerado, particularmente em partes da planta que possuem pouca contribuição para a produção ou a qualidade (HARDWICK, 2006).

Sistemas de previsão de doenças de plantas podem ser classificados de várias maneiras, de acordo com o tipo de informação usada para fazer a previsão ou com a abordagem conceitual para a previsão. Especificamente, previsores podem ser classificados com respeito a (CAMPBELL; MADDEN, 1990): se informação da cultura, da doença, do patógeno ou do ambiente, ou uma combinação dessas, são usadas para fazer as previsões; se as previsões são pré-plantio ou pós-plantio; se informação empírica ou fundamental foi usada no desenvolvimento do sistema; e se características específicas das epidemias, tais como inóculo primário, inóculo secundário ou taxa de aumento da doença, servem de base para as previsões. Uma ampla revisão bibliográfica de sistemas de previsão de doenças de plantas, abrangendo importantes doenças de várias culturas agrícolas, encontra-se em Reis e Bresolin (2004).

Um sistema de alerta, para que tenha sucesso, precisa ser adotado e implementado pelos produtores, devendo haver a percepção de que é possível obter benefícios específicos e tangíveis com o seu uso. Atributos que asseguram o sucesso incluem (CAMPBELL; MADDEN, 1990): confiabilidade, simplicidade para implementar, importância da doença, utilidade do alerta, disponibilidade aos produtores, aplicabilidade a várias doenças/pragas e eficiência de custo.

Segundo Bourke (1970), para desenvolver um sistema de alerta é preciso que a doença satisfaça quatro requisitos: a doença causa perdas economicamente significativas na qualidade ou na quantidade da produção; a doença varia entre cada estação de cultivo; medidas de controle da doença estão disponíveis e são economicamente viáveis; e há informação suficiente a respeito da natureza da dependência da doença em relação às condições meteorológicas.

Muitos estudos a respeito do papel das condições meteorológicas nas doenças de plantas indicam que a doença é mais afetada pelas condições microclimáticas no dossel das plantas do que pelas condições macroclimáticas medidas a uma certa distância da cultura em estações meteorológicas padrão. Entretanto, condições macroclimáticas produzem o microclima - é possível usar regras para determinar relacionamentos entre o macro e o microclima - e existe um limite na extensão com que o microclima pode facilitar o desenvolvimento da doença sob condições macroclimáticas desfavoráveis (COAKLEY, 1988). Ele chegou ainda a sugerir que tentativas de relacionar dados macroclimáticos com doenças podem ter alcançado resultados limitados, em parte, por causa da dificuldade de analisar grandes quantidades de dados sem o auxílio de computador.

Anteriormente, desvantagens de usar dados microclimáticos de dentro dos campos de cultivo no desenvolvimento e/ou implementação de sistemas de alerta incluíam períodos de tempo relativamente curtos e poucos locais para os quais estavam disponíveis os dados, alto custo de coleta desses dados e frequência considerável de perda de dados devido a falhas nos equipamentos (COAKLEY, 1988). Nos últimos anos, os instrumentos manuais deram lugar às estações meteorológicas automáticas, que podem ser instaladas próximas aos campos de cultivo. A cada geração, essas estações tornam-se mais sofisticadas, confiáveis e de menor preço. Mesmo assim, exigir de produtores a instalação e a manutenção dessas estações meteorológicas e o gerenciamento dos dados de cada campo não é prático nem economicamente viável.

Como alternativa a essa questão, esforços têm sido feitos para desenvolver e validar sistemas de alerta que utilizem dados de redes regionais de estações meteorológicas (GENT; SCHWARTZ, 2003) e tecnologias para a obtenção de dados de locais específicos sem sensores in loco (MAGAREY et al., 2001). Outra tendência é utilizar dados meteorológicos estimados a partir de modelos de previsão do tempo para antecipar ainda mais os alertas, permitindo um tempo

maior para a tomada de decisão e para a aplicação de medidas de controle (SHTIENBERG ; ELAD, 1997).

O foco principal desta seção é apresentar modelos de previsão e um sistema de alerta da ferrugem do cafeeiro desenvolvidos pela Embrapa Informática Agropecuária. Outra iniciativa semelhante da Embrapa é o SisAlert, um sistema web de previsão de doenças de plantas baseado em modelos de simulação modulares e genéricos para prever o estabelecimento de doenças a partir de dados meteorológicos obtidos de estações meteorológicas automáticas e de prognósticos de tempo de curto prazo (FERNANDES et. al., 2011). Esse sistema está operacional para duas doenças do trigo e cinco doenças da cultura da maçã, além de ter modelos de outras culturas agrícolas em avaliação e em desenvolvimento.

3.1 Modelos de previsão de doenças de plantas

Modelos representam a percepção (ou imaginação) da realidade de forma simbólica e simplificada. Modelagem e simplificação são essenciais ao processo científico. A solução ideal é o modelo contemplar os aspectos essenciais do sistema real pertinentes ao problema em questão. A complexidade ou simplicidade do modelo deve estar em acordo com o seu propósito - enquanto a simplicidade facilita o entendimento do modelo, a complexidade pode permitir maior acurácia na descrição do sistema (CAMPBELL et al., 1988).

Modelos podem ser classificados em dois grupos, dependendo da abordagem de desenvolvimento. O primeiro tipo de modelo é chamado fundamental ou mecânico. O desenvolvimento desses modelos parte de um conceito, hipótese ou teoria, em vez de um conjunto de dados. Um modelo consistente com o conceito é elaborado e depois são realizados experimentos para testar a sua acurácia (CAMPBELL; MADDEN, 1990).

Esses modelos são derivados de tentativas de compreensão da realidade, sendo que essa compreensão pode ser obtida a partir de experimentos prévios - em laboratório, câmara de ambiente controlado, casa de vegetação ou campo - ou de princípios biológicos (CAMPBELL et al., 1988).

Modelos de previsão fundamentais são geralmente simples e baseados em um ou poucos componentes do ciclo da doença, nos quais a infecção é o componente que prevalece. Contudo, é possível considerar o ciclo completo da doença por meio de uma abordagem de análise sistêmica e modelos de simulação (MADDEN; ELLIS, 1988).

O segundo tipo de modelo é chamado empírico ou correlativo. Os modelos empíricos são desenvolvidos a partir da coleta e análise de dados atuais e históricos sobre níveis da doença e outros fatores bióticos e abióticos (MADDEN; ELLIS, 1988). Eles descrevem um relacionamento observado entre duas ou mais variáveis do conjunto de dados, normalmente derivados a partir do ajuste dos dados a um modelo aceitável (CAMPBELL; MADDEN, 1990). Conhecimento teórico relacionado a mecanismos básicos não é exigido (CAMPBELL et al., 1988).

Modelos empíricos estão relacionados com uma previsão apenas por estação de cultivo ou podem envolver múltiplas previsões. Os primeiros são úteis quando é importante prever o inóculo inicial ou o nível inicial da doença, e o seu desenvolvimento geralmente requer observações de vários anos e/ou locais. Aqueles que envolvem múltiplas previsões são úteis quando a doença pode aumentar rapidamente durante a estação de crescimento, ou quando o valor econômico da cultura justifica várias intervenções de controle. O seu desenvolvimento requer observações feitas sobre a doença, o ambiente e/ou outros fatores por toda a estação de cultivo. Em geral, mais

de dois anos ou locais precisam ser estudados e anos adicionais são necessários para a validação (MADDEN; ELLIS, 1988).

Coakley (1988), baseado em sua experiência, sugeriu um mínimo de oito a doze anos de registro de dados, de campos com fontes naturais de inóculo, para identificar com segurança quais podem ser os fatores climáticos de influência no desenvolvimento de uma doença. Sugeriu também, quando se tem menos de oito anos de registro, que dados de diferentes localidades de uma região geográfica podem ser utilizados.

Madden e Ellis (1988) indicaram duas formas pelas quais modelos de previsão empíricos podem ser desenvolvidos. A primeira, chamada de qualitativa, envolve o desenvolvimento de critérios de previsão sem qualquer análise estatística formal. Alguns exemplos de modelos desenvolvidos dessa forma são: previsão da mancha preta do amendoim (JENSEN; BOYLE, 1966; PARVIN JUNIOR. et al., 1974; PEDRO JÚNIOR et al., 1994), previsão da requeima da batateira (COSTA et al., 2002; WALLIN, 1962), previsão da pinta-preta do tomateiro (MADDEN et al., 1978) e previsão da queima das folhas da cenoura (SOUZA et al., 2002).

A segunda forma de desenvolvimento dos modelos de previsão, chamada de quantitativa, é baseada em análise estatística e modelagem dos dados observados. Os métodos e técnicas utilizados são variados. Em uma retrospectiva dos modelos apresentados na literatura, é possível perceber que o emprego desses métodos e técnicas acompanha a evolução nas disciplinas relacionadas com a análise de dados.

O predomínio é de métodos e técnicas estatísticos, sendo a análise de regressão a mais popular (MADDEN; ELLIS, 1988). Como exemplo, cita-se o uso de regressão linear múltipla no desenvolvimento de modelos para prever a severidade de epidemias da ferrugem asiática da soja (DEL PONTE et al., 2006). Alguns métodos matemáticos e estatísticos para a modelagem de dados epidemiológicos foram revisados por Hau e Kranz (1990).

Mais recentemente, os trabalhos publicados relatam o uso de métodos e técnicas modernos e sofisticados, não necessariamente inovadores, mas que ganharam/recuperaram visibilidade e importância nos últimos anos, como redes neurais (PAUL; MUNKVOLD, 2005), regressão logística (DE WOLF et al., 2003), árvores de decisão (MEIRA et al., 2009) e análise estatística de séries temporais (XU et al., 2000).

3.2 Modelagem para previsão e alerta da ferrugem do cafeeiro

A ferrugem, causada pelo fungo *Hemileia vastatrix* Berk. & Br., é a principal doença do cafeeiro em todo o mundo. No Brasil, em regiões onde as condições climáticas são favoráveis à doença, os prejuízos na produção atingem cerca de 35% em média, podendo chegar a mais de 50%. Os principais danos causados pela ferrugem são a queda precoce das folhas e a seca dos ramos, que, em consequência, não produzem no ano seguinte (ZAMBOLIM et al., 2005).

Além da importância econômica, a ferrugem atende outros requisitos, como a variação na sua intensidade entre cada estação de cultivo e a disponibilidade de medidas de controle economicamente viáveis, que justificam o desenvolvimento de modelos de previsão e um sistema de alerta ou aviso.

O cafeeiro alterna anos de alta e baixa produção (carga pendente de frutos), caracterizando um ciclo bienal. A ferrugem ataca com maior intensidade em anos de alta carga. A curva de progres-

so padrão da doença inicia-se em dezembro/janeiro, aumenta em escala logarítmica de março a abril e atinge o pico por volta de junho. A partir daí, decresce devido às baixas temperaturas e à desfolha das plantas ocasionada pela colheita, pela senescência natural e também pela severidade da doença.

A elevação da temperatura e da precipitação pode atrasar o começo da epidemia e, simultaneamente, a ocorrência de chuvas esporádicas e aumento na temperatura média entre abril e julho permitem à doença manter altos índices de incidência até agosto. Nesses casos, torna-se necessário rever e readaptar as medidas de controle (CHALFOUN et al., 2001). O controle da ferrugem envolve principalmente o uso de fungicidas protetores e sistêmicos. O uso de variedades resistentes também é importante no controle da doença (ZAMBOLIM et al., 2005).

A modelagem da ferrugem do cafeeiro para estudos epidemiológicos é encontrada na literatura científica e existem exemplos de modelos de previsão da doença, a maioria deles empíricos. O ajuste dos dados observados a equações de regressão foi a técnica mais comum de modelagem, mas existem exemplos também de modelo fundamental e de modelo empírico com abordagem qualitativa. Trabalhos mais recentes utilizam técnicas de maior visibilidade no momento atual, como as consideradas da área de mineração de dados.

3.2.1 Previsão com modelo fundamental

Kushalappa e Eskes (1989) propuseram o desenvolvimento de um modelo compreensivo capaz de explicar o curso de ação biológica do patógeno e integrar os vários fatores que influenciam o sistema, argumentando que esses diversos fatores que afetam o progresso da ferrugem do cafeeiro no campo não poderiam ser identificados por um experimento ou por procedimentos estatísticos clássicos, como a análise de regressão.

O sistema epidêmico da ferrugem do cafeeiro é composto de processos epidemiológicos policíclicos, que consistem de uma série de processos monocíclicos. O princípio de uma epidemia começa com um inóculo inicial e cada ciclo da doença (processo monocíclico) é formado pelos macroprocessos de esporulação, disseminação e infecção. Todos constituem os componentes estruturais do sistema epidêmico (KUSHALAPPA, 1994).

Baseado nesses aspectos, foi desenvolvido um modelo de previsão da taxa de progresso da ferrugem, considerando o inóculo inicial e fatores significativos do ambiente e do hospedeiro que influenciam o processo monocíclico de *H. vastatrix*. Denominado de “razão de sobrevivência líquida para o processo monocíclico” (RSLPM), o modelo foi formado pela integração de modelos fundamentais e empíricos desenvolvidos para cada um dos componentes estruturais da doença (KUSHALAPPA et al., 1983, 1984).

Na obtenção do modelo RSLPM, vários fatores que influenciam o progresso da ferrugem do cafeeiro, relacionados com o hospedeiro, o patógeno e o ambiente – ou, especificamente, os micro e mesoprocessos componentes dos macroprocessos – foram transformados em “equivalentes de processo” para o ambiente e o hospedeiro. Depois, os produtos multiplicativos dos equivalentes de micro e mesoprocessos foram derivados, designados de “equivalentes de processo monocíclico” para o ambiente e para o hospedeiro. Estes e o nível de inóculo foram então transformados em outro parâmetro multiplicativo, a razão de sobrevivência líquida para o processo monocíclico de *H. vastatrix* (KUSHALAPPA, 1989a). A incorporação dos três componentes do triângulo de doenças de plantas no modelo foi baseada na atividade biológica do fungo, o que fez os criadores do modelo o considerarem do tipo fundamental.

A influência do inóculo, denominada de razão de sobrevivência básica (RSB), foi quantificada com base na incidência (proporção de folhas com ferrugem) ou na severidade (proporção de área foliar com ferrugem).

A influência do ambiente, ou equivalente de processo monocíclico para o ambiente (EPMA), foi calculada pela multiplicação de sua influência nos processos de disseminação e de infecção. O equivalente de disseminação foi determinado em função da velocidade diária do vento, da quantidade de chuva diária e da densidade de plantas, enquanto o equivalente de infecção foi determinado em função da duração do molhamento foliar (em horas) e da temperatura durante esse período.

O equivalente de processo monocíclico para o hospedeiro (EPMH) foi determinado pelo equivalente de processo devido à predisposição do hospedeiro ao ataque da ferrugem por causa de alta produção.

Por fim, a razão de sobrevivência líquida para processo monocíclico foi definida como o produto final da multiplicação das influências do inóculo, do ambiente e do hospedeiro, conforme a equação 1 (KUSHALAPPA, 1989a; KUSHALAPPA et al., 1983).

$$RSLPM = RSB \times EPMA \times EPMH \quad (1)$$

Valores de RSLPM a partir de dados observados no campo, para um intervalo de 28 dias antes da data de predição (DP), foram relacionados, por análise de regressão, com taxas de infecção da ferrugem observadas 28 dias após DP, corrigidas para o crescimento do hospedeiro (KUSHALAPPA et al., 1984). Esse intervalo de predição foi escolhido por causa da média observada do período latente do fungo, de outubro a março, que foi de aproximadamente 28 dias.

Várias equações foram desenvolvidas para predizer a taxa de infecção, considerando diferentes parâmetros de área de produção de inóculo no cálculo dos valores de RSLPM (KUSHALAPPA et al., 1984). As equações que obtiveram os melhores coeficientes de determinação (R^2) foram:

$$k'' = 0,00044 + 14,766 \times RSLPM - 2511,21 \times RSLPM^2 \quad (2)$$

$$k'' = 0,023 + 14,026 \times RSLPM - 87,382 \times RSLPM^2 \quad (3)$$

onde a equação 2 é para predizer a severidade da doença, considerando a proporção de área foliar com ferrugem como o parâmetro RSB; a equação 3 é para predizer a incidência da doença, considerando a proporção de folhas com ferrugem como o parâmetro RSB; k'' é a taxa de infecção para 28 dias após DP, corrigida para o crescimento do hospedeiro; RSLPM é dada pela média diária da razão de sobrevivência líquida para o processo monocíclico durante 28 dias antes de DP. As equações 2 e 3 explicaram 76% e 64% (R^2 igual a 0,76 e 0,64), respectivamente, da variação em k'' .

Considerando que uma incidência de ferrugem de cerca de 10% justificaria uma aplicação de fungicida, um limite do valor de RSLPM para recomendar aplicações de fungicida foi derivado

pela substituição de $k'' = 0,1$ na equação 3 (KUSHALAPPA, 1989b; KUSHALAPPA et al., 1984). Esse limite foi $RSLPM = 0,0057$, considerando a proporção de folhas com ferrugem como inóculo (RSB). Substituições semelhantes em outras equações permitiram se chegar no limite $RSLPM = 0,00015$, considerando a proporção de área foliar com ferrugem como inóculo.

A partir desses limites, um sistema de alerta simples e outro mais complexo foram desenvolvidos para recomendar aplicações de fungicida para o controle da ferrugem do cafeeiro. O sistema simples se resumiu em uma tabela, que foi formada pelo agrupamento dos valores de inóculo (RSB), de produção ($EPMH$) e de condições do ambiente ($EPMA$), observados no Estado de Minas Gerais, em certos intervalos convenientes.

No caso do sistema simples, em intervalos quinzenais, deve-se quantificar a incidência (percentual de folhas atacadas) ou a severidade (área foliar atacada) da ferrugem e indicar a produção como alta ou baixa; depois, basta consultar a tabela sobre a recomendação ou não de aplicação de fungicida.

No caso do sistema complexo, deve-se quantificar $RSLPM$, em intervalos de 14 dias, e se recomenda a aplicação de fungicida quando o limite preestabelecido é igualado ou superado – $RSLPM \geq 0,0057$ ou $RSLPM \geq 0,00015$, caso se tenha quantificado a incidência ou a severidade da doença, respectivamente. Ambos os sistemas foram validados em condições de campo e foram considerados eficientes na determinação das épocas oportunas de aplicação de fungicidas (KUSHALAPPA et al., 1986).

3.2.2 Previsão com modelo empírico pela abordagem qualitativa

Modificações no clima, nos últimos anos, têm ocasionado alterações na severidade da ferrugem, bem como no início e no pico da doença em algumas regiões do Brasil (ZAMBOLIM et al., 2002). Diante dessas alterações, surgiu a dúvida sobre a aplicação de fungicidas sistêmicos por meio de duas pulverizações foliares ou aplicações via solo, para que se obtivesse controle racional e econômico da ferrugem.

Para se identificar, então, os períodos favoráveis à ferrugem, nos quais as plantas deveriam ser atomizadas, foi desenvolvido um sistema de previsão ou de aviso (GARÇON et al., 2004). O objetivo foi criar um sistema simples - simplicidade é um atributo importante para a aceitação do sistema, pois maiores são as chances de adoção pelos agricultores - e confiável de prever o desenvolvimento da doença no campo, determinando o momento propício para iniciar o controle químico por meio de pulverizações com fungicida sistêmico, bem como o intervalo entre as aplicações.

As variáveis meteorológicas empregadas no sistema de previsão foram o molhamento foliar diário e a temperatura média durante esse período de molhamento, obtidas em estação meteorológica colocada no meio da área experimental. Com os dados diários dessas variáveis meteorológicas calculou-se o valor de severidade da ferrugem (VSF), a partir de uma matriz de valores de severidade semelhante à idealizada por Wallin (1962) para a requeima da batateira, modificada para a ferrugem do cafeeiro (Tabela 1).

O limiar de ação para indicação do momento da pulverização foi baseado no acúmulo dos valores diários de VSF. Os limites de VSF estipulados para o teste e a validação do sistema foram 29, 34, 39 e 44, para anos de alta carga pendente de frutos (alta intensidade da ferrugem), e 49, 59, 69 e 79, para anos de baixa ou média carga pendente (baixa intensidade da ferrugem).

Tabela 1. Matriz para cálculo dos valores de severidade da ferrugem (VSF) do cafeeiro, com base no período de molhamento foliar e na temperatura média do período.

Molhamento foliar (h/diárias)	Temperatura (°C)						
	< 16	16-18	19-20	21-24	25-26	27-29	30
0	0*	0	0	0	0	0	0
0 < h ≤ 8	0	0	1	2	1	0	0
8 < h ≤ 17	0	1	2	3	2	1	0
17 < h ≤ 24	0	2	3	4	3	2	0
h = 24**	0	0	1	2	1	0	0

* Valor de severidade da ferrugem (VSF) diário.

** Molhamento foliar diário de 24 h, porém este dentro de um período de molhamento de mais de 48 h sem interrupção.

Obs.: se dentro de 30 dias não houver acumulado mais de 5 VSF, desconsideram os VSF acumulados até o momento.

Fonte: Garçon et al. (2004).

Em uma lavoura de alta carga pendente (101,5 sacas beneficiadas/hectare), foram recomendadas duas pulverizações com fungicida sistêmico quando o valor acumulado de VSF atingiu 29-31, igualando-se as duas aplicações do tratamento com calendário fixo. Em uma lavoura de média carga pendente (22,4 sacas beneficiadas/hectare), recomendou-se uma única pulverização quando o valor acumulado de VSF atingiu 49-51, enquanto todos os outros tratamentos demandaram duas aplicações. Portanto, o sistema baseado no VSF foi tão eficiente quanto o calendário no controle da ferrugem do cafeeiro, porém com economia de uma pulverização na lavoura com carga média de frutos.

Os resultados mostraram que, normalmente, em lavouras com baixa a média carga pendente de frutos, uma única aplicação de fungicida sistêmico, no momento oportuno, indicado com base no número de horas de molhamento foliar e na temperatura média durante o período de molhamento, como sugerido pelo modelo de aviso proposto, foi suficiente para se alcançar um controle eficiente e racional da ferrugem do cafeeiro.

3.2.3 Previsão com modelos empíricos pela abordagem quantitativa

A modelagem da ferrugem do cafeeiro para estudos epidemiológicos encontrada na literatura científica revela diferentes exemplos de modelos de previsão da doença, a maioria deles empíricos desenvolvidos segundo a abordagem quantitativa. O ajuste dos dados observados a equações de regressão é a técnica mais comum de modelagem utilizada (KUSHALAPPA; ESKES, 1989; MORAES et al., 1976; ZAMBOLIM et al., 2002).

Moraes et al. (1976) observaram que o período de incubação (PI) - período decorrido da inoculação até a formação de 50% de pústulas - tendeu a encurtar nos meses mais quentes (28 dias) e tornar-se mais longo nos meses mais frios (65 dias). Os autores sugeriram a utilização da seguinte equação para a estimativa do PI:

$$y = 103,01 - 0,98 \times x_1 - 2,1 \times x_2 \quad (4)$$

onde y é a estimativa do período de incubação em dias, x_1 é a temperatura média máxima e x_2 a temperatura média mínima durante o período.

Essa equação foi usada para dar uma ideia de quão severa a ferrugem poderia ser durante certas estações ou meses do ano. Moraes (1983) sugeriu, com base em estimativas obtidas pela equação 4, para o período de outubro a março, os seguintes níveis de severidade de ataque da ferrugem do cafeeiro: risco alto de ataque severo, quando o PI estimado for inferior a 35 dias; risco médio de ataque severo quando o PI for estimado entre 35 e 45 dias; e pequena probabilidade de risco de ataque severo, quando o PI estimado for superior a 45 dias.

Chaves et al. (1970) citado por Kushalappa (1989b), logo que a ferrugem do cafeeiro surgiu no Brasil, consideraram desnecessárias aplicações de fungicida, de maio a agosto, devido aos períodos latentes mais longos. Kushalappa (1989b) argumentou que a razão de não haver necessidade de aplicações de fungicida nos meses mais frios seria diferente: a temperatura durante o período de molhamento foliar geralmente fica abaixo de 15°C, o que é limitante para a infecção.

Alfonsi et al. (1974) estudaram a associação entre níveis de infecção (média de pústulas por folha), variáveis climáticas e área foliar das plantas. A média das temperaturas máximas, a média das temperaturas mínimas e o total de chuvas, registrados em períodos de 15, 30 e 45 dias, foram correlacionados com os níveis de infecção observados ao final dos respectivos períodos. Os coeficientes de determinação (R^2) obtidos entre o nível de infecção e as três variáveis climáticas, independente da área foliar, mostraram que a associação com o período de 45 dias expressou melhor a proporção de acréscimos de pústulas (cerca de 95% de explicação da variação na severidade da doença). Na literatura consultada, não há registro de sistema de alerta da ferrugem do cafeeiro baseado nas equações de regressão obtidas.

Também por análise de regressão, vários fatores meteorológicos e biológicos foram considerados para explicar a taxa de progresso da ferrugem (KUSHALAPPA; ESKES, 1989). Os fatores mais significativos foram identificados pelo critério de seleção stepwise. Como variáveis dependentes, foram consideradas a severidade da doença na data de previsão (DP) e a taxa de infecção da ferrugem para os intervalos de um a dois períodos latentes (28 dias) após DP. A equação que explicou a máxima variação (94%) na taxa de infecção foi:

$$k'' = 0,031 + 4,881 \times PAFE + 0,022 \times PNF - 0,001 \times MIN - 0,001 \times MAX - 0,001 \times CHUVA \quad (5)$$

onde k'' é a taxa de infecção, corrigida para o crescimento do hospedeiro, para 56 dias depois de DP; *PAFE* é a proporção de área foliar com esporos na DP; *PNF* é a proporção de novas folhas formadas durante 14 dias antes de DP; *MIN* é a média das mínimas e *MAX* é a média das máximas temperaturas (em °C) para 14 dias antes de DP; *CHUVA* é o total de chuvas (em mm) entre 14 a 28 dias antes de DP. Não foi encontrado, na revisão bibliográfica, trabalho a respeito do uso das equações desenvolvidas em sistema de alerta da ferrugem do cafeeiro.

Correlações significativas foram observadas entre variáveis independentes usadas na formulação de equações de regressão para prever a taxa de progresso da ferrugem do cafeeiro (KUSHALAPPA et al., 1983). Como resultado, alguns parâmetros, que independentemente explicaram variação significativa na doença, foram eliminados devido à multicolinearidade. O sucesso preditivo de tais modelos depende da ocorrência futura dos diferentes parâmetros, incluindo aqueles não usados no modelo, em combinações semelhantes àquelas observadas. Isso

torna esse tipo de modelo menos estável sob condições de campo, a menos que esteja baseado em vários anos de registro de dados (KUSHALAPPA; ESKES, 1989).

Mais recente, Pinto et al. (2002) avaliaram o potencial de redes neurais para descrever a epidemia da ferrugem do cafeeiro. Elas foram empregadas para estabelecer relações entre variáveis climáticas e produção e a incidência da ferrugem. As seguintes variáveis foram utilizadas para construir as redes: precipitação pluvial; número de dias com e sem precipitação; umidade relativa do ar; horas de insolação; temperaturas média, máxima e mínima, calculadas como médias ou somatórios para os 15, 30, 45 e 60 dias anteriores às avaliações da incidência da ferrugem; e a variável produção, a qual assumiu valor '0' para as plantas antes do início da produção e '1' para as plantas em fase de produção. Séries temporais da incidência da doença, isoladamente, também foram utilizadas na elaboração de redes neurais.

A camada de entrada para as redes foi formada pelas variáveis climáticas mais a variável de produção (variáveis independentes) e a variável de saída foi a incidência da ferrugem (variável dependente). Os menores valores do erro médio de previsão (EMP = 1,17%) e do quadrado médio do desvio (QMD = 3,43) foram obtidos para a rede neural elaborada com as variáveis produção, umidade relativa, horas de insolação e temperatura mínima, relativas ao período de 30 dias anteriores à avaliação da incidência da doença. A melhor rede neural (EMP = 4,72% e QMD = 3,95) elaborada a partir das séries temporais teve como variáveis de entrada as observações da incidência da doença de quatro quinzenas anteriores à data de avaliação (PINTO et al., 2002).

As redes neurais elaboradas poderiam ser utilizadas como modelos de previsão da ferrugem do cafeeiro. Apesar do melhor desempenho do modelo que incluiu as variáveis climáticas, os autores sugeriram que o uso de séries temporais, apenas, poderia facilitar a previsão de epidemias da doença. Eles consideraram que avaliar a intensidade da doença é mais fácil para o produtor ou o agente de extensão do que coletar variáveis climáticas.

A indução de árvores de decisão é uma técnica de modelagem alternativa. O seu propósito é descobrir a estrutura preditiva do problema e/ou produzir modelos de predição precisos (BREIMAN et al., 1984). A indução não é prejudicada por multicolinearidade entre as variáveis independentes como na regressão (BUTT; ROYLE, 1990). Além disso, diversas variáveis, numéricas ou categóricas, podem ser analisadas ao mesmo tempo, sendo que o próprio algoritmo de indução se encarrega de selecionar as de maior importância.

A árvore de decisão é um modelo representado graficamente por nós e ramos, parecido com uma árvore, mas no sentido invertido (WITTEN et al., 2011). O nó raiz, no topo da estrutura, e os nós internos são nós de decisão. Cada um contém um teste sobre uma variável independente e os resultados desse teste formam os ramos da árvore. Os nós folhas, nas extremidades, representam valores de predição da variável dependente ou distribuições de probabilidade desses valores.

Paul e Munkvold (2004) usaram este tipo de modelagem para prever categorias de severidade da cercosporiose do milho em estágio avançado do cultivo. Árvores de decisão também modelaram epidemias de giberela do trigo, procurando prever se a severidade da doença seria maior ou igual a 10% (MOLINEROS et al., 2005). Baker et al. (1993) desenvolveram uma árvore de decisão para prever o risco (alto ou baixo) de mortalidade de pinus em decorrência de podridão das raízes causada por *Heterobasidion annosum*.

Com relação à ferrugem do cafeeiro, Meira et al. (2008) analisaram manifestações epidêmicas da doença com o auxílio de uma árvore de decisão. As variáveis explicativas mais importantes foram a temperatura média nos períodos de molhamento foliar, a carga pendente de frutos, a média das temperaturas máximas diárias no período de incubação e a umidade relativa do ar. A árvore de decisão demonstrou seu potencial simbólico e interpretável, auxiliando na compreensão de quais variáveis e como as interações dessas variáveis conduziram as epidemias da ferrugem no campo.

Modelos em árvore de decisão foram desenvolvidos para alerta da ferrugem do cafeeiro em lavouras com alta carga pendente de frutos e em lavouras com baixa carga (MEIRA, 2008; MEIRA; RODRIGUES, 2009; MEIRA et al., 2009). Dados de incidência mensal da doença no campo coletados durante oito anos foram transformados em valores binários considerando limites de 5 e 10 pontos percentuais (p.p.) na taxa de infecção. O alerta é indicado quando a taxa de infecção, prevista para o prazo de um mês, atingir ou ultrapassar o respectivo limite.

A acurácia do modelo para lavouras com alta carga pendente e limite de 5 p.p. foi de 81%, calculada por validação cruzada, também com valores bons e equilibrados para outras medidas de avaliação; o modelo para o limite de 10 p.p. teve acurácia de 79%, sem apresentar o mesmo equilíbrio entre as demais medidas. Os modelos para lavouras com baixa carga pendente não apresentaram o mesmo bom desempenho.

A partir do mesmo conjunto de dados preparados por Meira (2008), Cintra et al. (2011) desenvolveram modelos com um método de indução de árvores de decisão fuzzy. O método, denominado Fuzzy DT, é baseado no algoritmo clássico C 4.5 (QUINLAN, 1993), acrescentado de características interessantes da lógica fuzzy relacionadas à interpretabilidade dos modelos e à manipulação de atributos contínuos.

Valores numéricos dos atributos preditivos foram transformados em três valores linguísticos fuzzy: baixo, médio e alto. Foram utilizados conjuntos fuzzy triangulares uniformemente distribuídos nas três partições. O número de conjuntos fuzzy foi definido empiricamente. Os experimentos utilizaram o método de raciocínio fuzzy tradicional, que classifica um novo exemplo usando a classe da regra com a maior compatibilidade com ele (CINTRA et al., 2011).

Comparados com modelos correspondentes obtidos pelo método clássico, os modelos fuzzy apresentaram taxas de erro competitivas (CINTRA et al., 2011). Além disso, o uso dos termos linguísticos evitaram alterações bruscas de decisão com base em limites numéricos dos atributos preditivos, uma desvantagem do método tradicional, e permitiram uma interpretação mais natural das regras dos modelos.

Foram ainda desenvolvidos, em colaboração com grupos de pesquisa internacionais, modelos para a ferrugem do cafeeiro com outras técnicas de modelagem: equações de regressão capazes de prever intervalos em vez de pontos, usando support vector machines (SVM) (LUACES et al., 2010); modelos de predição obtidos por aprendizado não determinístico (LUACES et al., 2011); e predição usando redes bayesianas (PÉREZ-ARIZA et al., 2012).

Modelos selecionados foram validados com dados diferentes dos usados no treinamento desses modelos (DI GIROLAMO NETO et al., 2012). Uma vez que os modelos não mantiveram o bom desempenho nessa validação, iniciou-se um novo ciclo do processo de mineração de dados. O objetivo foi a evolução de modelos em árvore de decisão e a geração de novos modelos com

as técnicas máquinas de vetores suporte, redes neurais e florestas aleatórias (DI GIROLAMO NETO et al., 2014).

No novo ciclo do processo, além do acréscimo das técnicas de mineração de dados, houve outras mudanças em termos da metodologia (DI GIROLAMO NETO et al., 2014): inclusão de dados de dois outros municípios da região cafeeira Sul de Minas, com características climáticas distintas; uso de método de balanceamento de classes nos dados oriundos das lavouras de baixa carga pendente; seleção de atributos por meio de métodos objetivos e automáticos; e utilização do espaço Receiver Operating Characteristic (ROC) para avaliação, comparação e seleção dos melhores modelos.

As técnicas de máquinas de vetores suporte e de florestas aleatórias geraram os modelos com os melhores desempenhos. O procedimento de balanceamento de classes melhorou a taxa de acerto dos modelos para lavouras com baixa carga pendente. Para alta e baixa cargas pendentes, as melhores taxas de acerto foram 85,3% e 88,9%, respectivamente. Outras medidas de desempenho como sensibilidade e especificidade também apresentaram valores altos e equilibrados (DI GIROLAMO NETO et al., 2014).

3.3 Sistema de alerta da ferrugem do cafeeiro

Sistemas de alerta que utilizam modelos de previsão ajudam a promover o uso racional de agroquímicos, ao indicar as condições que favorecem ou deixam de favorecer uma doença ou praga, permitindo agir somente quando necessário. A dificuldade de adoção desses sistemas é um problema comum, que inclui certas restrições (CAMPBELL; MADDEN, 1990): complexidade do modelo; dificuldade de obtenção dos dados necessários ao uso do modelo; e custos de implantação e manutenção do sistema para os produtores.

A abordagem normalmente utilizada no desenvolvimento de modelos empíricos consiste em realizar experimentos de curta duração (dois a quatro anos), em que equipamentos com sensores para a obtenção de dados meteorológicos são instalados no interior da cultura. Dessa forma, os modelos obtidos carecem de representatividade quanto às condições meteorológicas, pelo pouco tempo de observação dos dados, e o uso deles provavelmente vai requerer dos produtores a instalação e manutenção de equipamentos em suas lavouras.

Outro fator que contribui para a não adoção dos modelos é a falta de planejamento das atividades seguintes à modelagem. Muitas vezes, o próprio desenvolvimento do modelo é o principal objetivo, como forma de se compreender a epidemia da doença, terminando com uma avaliação estatística do seu desempenho. Não é planejada a sua validação, nem como colocá-lo em situação de uso real.

A Embrapa Informática Agropecuária, buscando contribuir para reduzir essas dificuldades, coordenou um projeto em que modelos de predição da ferrugem foram desenvolvidos com diferentes técnicas de mineração de dados (DI GIROLAMO NETO et al., 2014), a partir de uma infraestrutura estabelecida há vários anos de monitoramento de doenças e pragas do cafeeiro. Alguns dos modelos de melhor desempenho foram selecionados e incorporados em um sistema de alerta na Web (THAMADA et al., 2013). O propósito do sistema é servir como instrumento adicional de apoio aos técnicos da Fundação Procafé na elaboração dos comentários e das recomendações fitossanitárias para o público-alvo dos seus boletins de avisos mensais.

A Fundação Procafé, sediada em Varginha-MG, inaugurou em 1998 uma estação de avisos fitossanitários para a cultura do café. Com o passar dos anos, junto com o Ministério da Agricultura, Pecuária e Abastecimento (Mapa) e outras instituições parceiras, expandiram as estações de avisos para outros municípios da região Sul de Minas e para outras regiões produtoras do estado, como o Triângulo Mineiro e o Alto Paranaíba. Mais recentemente, iniciou também atuação na região Mogiana do Estado de São Paulo.

Em cada estação de avisos, são feitas avaliações mensais da incidência de doenças e pragas do cafeeiro e são registrados dados meteorológicos por meio de estação meteorológica automática. A partir desse monitoramento, são elaborados e divulgados avisos fitossanitários em boletins mensais destinados a técnicos de assistência ao produtor, imprensa, cafeicultores, estudantes e pesquisadores.

O Sistema de Alerta da Ferrugem do Cafeeiro - SafCafe (THAMADA et al., 2013), desenvolvido pela Embrapa com o apoio do Consórcio Pesquisa Café e em parceria com a Fundação Procafé, analisa e informa o risco de epidemias da doença por meio de modelos de alerta desenvolvidos com técnicas de mineração de dados, no âmbito das estações de avisos fitossanitários da Fundação Procafé na região Sul de Minas Gerais. Os modelos incorporados no sistema obtiveram desempenho satisfatório durante a modelagem, com acurácia em torno de 83% obtida por validação cruzada.

O sistema web SafCafe é carregado com arquivos de registros horários das estações meteorológicas instaladas nas cidades de Varginha, Carmo de Minas e Boa Esperança e prepara os dados de entrada para os modelos preditivos. Esses modelos avaliam as condições propícias ou desfavoráveis ao desenvolvimento da ferrugem, a partir dos dados preparados e de acordo com certas características das lavouras, como a carga pendente de frutos (alta ou baixa) e o tipo de plantio (largo ou adensado).

Como resultado, o SafCafe avisa para o mês seguinte sobre a possibilidade ou não de aumentos além de 5 ou 10 p.p. (pontos percentuais) na taxa de progresso da incidência da doença. A saída dos modelos é binária, emitindo alerta (1) quando é provável que, no próximo mês, a taxa de progresso da ferrugem seja maior ou igual ao limiar considerado. O não alerta (0) indica probabilidade alta de que a taxa não ultrapasse esse limiar.

No fluxo normal de uso do sistema de alerta, o técnico carrega, no início de cada mês, um arquivo contendo dados meteorológicos registrados ao longo do mês anterior e os valores do nível de incidência da ferrugem nas lavouras avaliados no final daquele mês. Os dados de incidência da doença não são usados na predição, mas são importantes para o acompanhamento da evolução da ferrugem no campo e do desempenho do sistema. Automaticamente após o carregamento, o SafCafe processa os registros para que possam ser utilizados pelos modelos e o sistema fica apto a realizar as predições.

Para realizar a predição, é necessário escolher como opções: a estação (Varginha, Carmo de Minas ou Boa Esperança); a carga pendente de frutos do cafeeiro (alta, baixa ou todas), o limiar de pontos percentuais da taxa de progresso da ferrugem a ser considerado (5 p.p., 10 p.p. ou todas); e o mês para qual é desejado realizar a predição. Escolhidas essas opções, o processo preditivo pode ser acionado.

A Figura 8 mostra uma predição realizada para Varginha, considerando todas as cargas e limiares da taxa de progresso, para agosto de 2013. As abas situadas logo abaixo da seleção das opções de

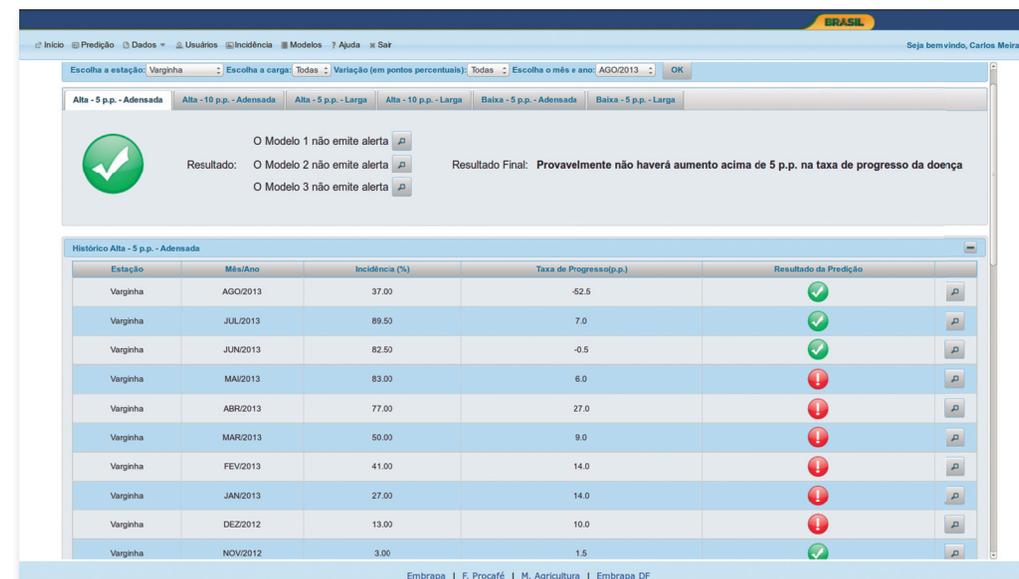


Figura 8. Tela de predição do SafCafe para agosto de 2013 em Varginha; destaque da aba de carga alta, lavoura adensada e limiar de 5p.p. da taxa de progresso da ferrugem do cafeeiro.

predição representam as combinações dessas opções. A aba selecionada (Alta - 5 p.p. - Adensada) mostra que o sistema indicou que não haveria aumento acima de 5 p.p. na taxa de progresso da ferrugem para aquele mês, considerando lavoura de carga alta e plantio adensado. O sistema exibe esse resultado textualmente e destaca-o por meio de um sinal verde (ícone maior verde à esquerda na Figura 8).

O resultado da predição do SafCafe considera um sistema de votação simples entre três modelos escolhidos e incorporados ao sistema. O alerta é positivo (sinalizado em vermelho) caso dois ou mais modelos indiquem esse resultado; o mesmo é feito para a emissão do não alerta (sinalizado em verde). No caso da Figura 8, os três modelos indicaram um não alerta para agosto de 2013 (Alta - 5 p.p. - Adensada).

Abaixo do resultado da predição para o mês escolhido, o sistema apresenta o histórico das predições nos meses anteriores, desde o início do ano agrícola (quadro inferior da tela do SafCafe na Figura 8). Além das predições anteriores, o sistema apresenta os valores de incidência e da taxa de progresso da doença que foram observados no campo, na realidade, permitindo, assim, acompanhar o desempenho do sistema com relação à acurácia (taxa de acerto) de seus alertas positivos e negativos.

Com a finalidade de adquirir a confiança necessária para que o SafCafe fosse utilizado no procedimento de elaboração dos boletins, foi realizada uma etapa de validação e confirmação do bom desempenho dos modelos incorporados ao sistema. O teste do sistema e dos seus modelos foi realizado com dados não utilizados no treinamento desses modelos durante a fase de modelagem. A avaliação foi realizada no período de três anos agrícolas (2011/2012, 2012/2013 e 2013/2014) com dados obtidos na Fazenda Experimental de Varginha.

A Tabela 2 apresenta os valores de acurácia do SafCafe nos anos agrícolas de validação, para as combinações de produtividade (carga alta e carga baixa) e tipo de plantio (adensado e

Tabela 2. Desempenho do SafCafé - Sistema de Alerta da Ferrugem do Cafeeiro na validação.

		Acurácia do sistema SafCafé (%)		
		Lavoura carga alta		Lavoura carga baixa
		Limiar 5 p.p.*	Limiar 10 p.p.*	Limiar 5 p.p.*
Plantio adensado	2011/2012	81,8	90,9	72,7
	2012/2013	90,9	90,9	81,8
	2013/2014	72,7	63,6	63,6
Plantio largo	2011/2012	72,7	90,9	63,6
	2012/2013	100	81,8	81,8
	2013/2014	81,8	63,6	63,6
Média	2011/2012	77,3	90,9	68,2
	2012/2013	95,5	86,4	81,8
	2013/2014	77,3	63,6	63,6
	2011 a 2014	83,3	80,3	71,2

* Alerta emitido quando a taxa de progresso da ferrugem do cafeeiro está prevista para aumentar além do limiar definido em pontos percentuais (5 p.p. ou 10 p.p.).

largo), como também as médias em cada ano e a média geral obtida no período todo de 2011 a 2014.

A acurácia do sistema no ano 2011/2012 ficou bem próxima da obtida no treinamento dos modelos, considerando carga alta e alerta para o limiar de 5 p.p. no plantio adensado (81,8%). No caso do plantio largo, houve um decréscimo da taxa de acerto (72,7%) devido a dois falsos positivos - previsão de aumento maior ou igual a 5 p.p. que não se confirmou no campo - emitidos pelo sistema. Também para carga alta, o desempenho em 2011/2012 dos alertas para o limiar de 10 p.p. foi ótimo (90,9%), nos dois tipos de plantio.

O SafCafé obteve os melhores resultados no ano agrícola 2012/2013, com desempenho excelente para carga alta e alerta para o limiar de 5 p.p. (90,9% no plantio adensado e 100% no plantio largo) e resultados muito bons do alerta para o limiar de 10 p.p. (90,9% e 81,8%, respectivamente). Em outras palavras, o sistema previu com exatidão a evolução da taxa de progresso da ferrugem, considerando o limiar de 5 p.p., durante todo o ano agrícola de carga alta na lavoura de plantio largo.

No ano agrícola 2013/2014, o desempenho do sistema não foi tão bom para as lavouras de carga alta, principalmente com relação ao alerta para o limiar de 10 p.p. (63,6%). O sistema acompanhou razoavelmente a evolução da ferrugem com relação aos alertas para o limiar de 5 p.p. (72,7% no plantio adensado e 81,8% no plantio largo), incluindo o período em que a doença esteve pressionada pela falta de chuva (janeiro a março), mas não identificou a evolução tardia da doença a partir de abril. Nos três meses em que a taxa de progresso da ferrugem no campo ultrapassou 10 p.p. (abril, junho e julho), não foram emitidos alertas considerando esse limiar.

A acurácia do sistema para lavouras de carga baixa foi boa somente no ano 2012/2013 (81,8% nos dois tipos de plantio). Nos outros anos avaliados, o desempenho foi fraco. Uma possível explicação é que o limiar do alerta em 5 p.p. pode não estar adequado, uma vez que a evolução da ferrugem é menos acelerada em anos de carga baixa. Sugere-se, então, a hipótese de que um

limiar menor do que 5 p.p. pode permitir um melhor desempenho dos modelos de previsão e, conseqüentemente, do sistema de alerta em anos de baixa carga pendente.

Um aspecto bastante importante a ressaltar é que o sistema de alerta foi exato na identificação do início da epidemia da ferrugem nos três anos agrícolas avaliados. No ano 2011/2012, o sistema acertou o início da epidemia em janeiro de 2012, indicando acertadamente que a evolução mais acentuada da doença não começaria em dezembro de 2011. Nos anos seguintes, 2012/2013 e 2013/2014, o sistema indicou corretamente o início das epidemias em dezembro. A confiabilidade desse tipo de informação é importante para ajudar o produtor a decidir pelo melhor momento de iniciar o controle da ferrugem do cafeeiro nas lavouras.

Conclui-se que o sistema de alerta da ferrugem do cafeeiro apresentou desempenho bastante satisfatório nos anos agrícolas avaliados, especialmente para lavouras com carga pendente alta e considerando os alertas para o limiar de 5 p.p. Dessa forma, acredita-se que o sistema SafCafé vai ser útil aos responsáveis pela elaboração dos comentários e das recomendações divulgados nos boletins de avisos fitossanitários da Fundação Procafé.

4 Considerações finais

Este capítulo mostrou como a aplicação de uma série de tecnologias pode contribuir significativamente para a redução de perdas e riscos associados a problemas sanitários que afetam virtualmente todas as áreas da agropecuária. Apesar dos avanços significativos já alcançados, é importante observar que a história da aplicação das TIC a problemas zoofitossanitários é muito recente, havendo ainda muitas oportunidades de pesquisa a serem exploradas no futuro. Com a evolução dos recursos computacionais disponíveis, tecnologias com custo computacional e financeiro proibitivos passam a ser opções viáveis, ampliando ainda mais as possibilidades de avanços.

Esforços de pesquisa futuros deverão se concentrar na criação e aplicação de novas tecnologias aos (muitos) problemas para os quais ainda não há uma solução satisfatória e estender tecnologias bem-sucedidas já desenvolvidas a problemas similares que ocorram em outras lavouras, criadouros e rebanhos. É importante enfatizar que muitas dessas tecnologias estão sendo planejadas, desenvolvidas e adaptadas de modo a explorar a crescente popularização e o aumento do poder computacional de dispositivos móveis, expandindo significativamente a aplicabilidade e o alcance das novas ferramentas desenvolvidas.

5 Referências

ALFONSI, R. R.; ORTOLANI, A. A.; PINTO, H. S.; PEDRO JUNIOR, M. J.; BRUNINI, O. Associação entre nível de infecção da ferrugem do cafeeiro, variáveis climáticas e área foliar, observadas em *Coffea arabica* L. In: CONGRESSO BRASILEIRO DE PESQUISAS CAFEEIRAS, 2., 1974, Poços de Caldas. **Resumos...** Rio de Janeiro: IBC-GERCA, 1974. p. 80-83.

BAKER, F. A.; VERBYLA, D. L.; HODGES, C. S.; ROSS, E. W. Classification and regression tree analysis for assessing hazard of pine mortality caused by *Heterobasidion annosum*. **Plant Disease**, St Paul, v. 77, n. 2, p. 136-139, 1993.

- BARBEDO, J. G. A. Automatically measuring early and late leaf spot lesions in peanut plants using digital image processing. In: CONGRESSO BRASILEIRO DE AGROINFORMÁTICA, 9., 2013, Cuiabá. **Agroinformática: inovação para a sustentabilidade do agronegócio brasileiro: anais**. Cuiabá: Universidade Federal de Mato Grosso, 2013a. Não paginado. SBIAgro 2013.
- BARBEDO, J. G. A. Automatic method for counting and measuring whiteflies in soybean leaves using digital image processing. In: CONGRESSO BRASILEIRO DE AGROINFORMÁTICA, 9., 2013, Cuiabá. **Agroinformática: inovação para a sustentabilidade do agronegócio brasileiro: anais**. Cuiabá: Universidade Federal de Mato Grosso, 2013b. Não paginado. SBIAgro 2013.
- BARBEDO, J. G. A. Automatic object counting in Neubauer chambers. In: SIMPÓSIO BRASILEIRO DE TELECOMUNICAÇÕES, 31., 2013, Fortaleza. [Anais...]. Rio de Janeiro: Sociedade Brasileira de Telecomunicações, 2013c. Não paginado.
- BARBEDO, J. G. A. An algorithm for counting microorganisms in digital images. **Revista IEEE América Latina**, v. 11, p. 1354-1359, 2013d.
- BARBEDO, J. G. A. A Study on the use of semi - automatic systems for counting objects in digital images. **International Journal of Computer and Information Technology**, v. 2, p. 1022-1028, Nov. 2013e.
- BARBEDO, J. G. A. **Digipathos**. 2013f. Disponível em: <<https://www.agropediabrasilis.cnptia.embrapa.br/web/digipathos>>. Acesso em: 2 set. 2014.
- BARBEDO, J. G. A. Computer-aided disease diagnosis in aquaculture: current state and perspectives for the future. **Revista Innover**, v. 1, n. 1, p. 19-32, 2014a.
- BARBEDO, J. G. A. An automatic method to detect and measure leaf disease symptoms using digital image processing. **Plant Disease**, St Paul, 2014b. Em fase de elaboração.
- BARBEDO, J. G. A. Using digital image processing for counting whiteflies on soybean leaves. **Journal of Asia-Pacific Entomology**, Taiwan, v. 17, p. 685-694, 2014c.
- BARBEDO, J. G. A. Method for counting microorganisms and colonies in microscopic images. In: INTERNATIONAL CONFERENCE ON COMPUTATIONAL SCIENCE AND ITS APPLICATIONS, 12., 2012, Salvador. **Proceedings...** Piscataway, NJ : IEEE, 2012. p. 83-87. ICCSA 2012.
- BOCK, C. H.; POOLE, G. H.; PARKER, P. E.; GOTTWALD, T. R. Plant disease severity estimated visually, by digital photography and image analysis, and by hyperspectral imaging. **Critical Reviews in Plant Sciences**, Boca Raton, v. 29, n. 1-3, p. 59-107, Jan./Jun. 2010. DOI: 10.1080/07352681003617285.
- BOISSARD, P.; MARTIN, V.; MOISAN, S. A cognitive vision approach to early pest detection in greenhouse crops. **Computers and Electronics in Agriculture**, Amsterdam, v. 62, n. 2, p. 81-93, July. 2008. DOI: 10.1016/j.compag.2007.11.009.
- BOURKE, P. M. A. Use of weather information in the prediction of plant disease epiphytotic. **Annual Review of Phytopathology**, Palo Alto, v. 8, n. 1, p. 345-370, Sept. 1970.
- BREIMAN, L.; FRIEDMAN, J. H.; OLSHEN, R. A.; STONE, C. J. **Classification and regression trees**. Boca Raton: CRC Press, 1984. 358 p.
- BUTT, D. J.; ROYLE, D. J. Multiple regression analysis in the epidemiology of plant diseases. In: KRANZ, J. (Ed.) **Epidemics of plant diseases: mathematical analysis and modeling**. 2nd ed. Berlin: Springer-Verlag, 1990. p. 143180.
- CAMARGO, A.; SMITH, J. S. Image pattern classification for the identification of disease causing agents in plants. **Computers and Electronics in Agriculture**, Nwe York, v. 66, n. 1, p. 121-125, May, 2009. DOI: 10.1016/j.compag.2009.01.003.
- CAMPBELL, C. L.; MADDEN, L. V. **Introduction to plant disease epidemiology**. New York: J. Wiley, 1990. 532 p.

- CAMPBELL, C. L.; REYNOLDS, K. M.; MADDEN, L. V. Modeling epidemics of root diseases and development of simulators. In: KRANZ, J.; ROTEM, J. (Ed.) **Experimental techniques in plant disease epidemiology**. Berlin: Springer-Verlag, 1988. p. 253-265.
- CHALFOUN, S. M.; CARVALHO, V. L. de; PEREIRA, M. C. Efeito de alterações climáticas sobre o progresso da ferrugem (*Hemileia vastatrix* Berk. & Br.) do cafeeiro (*Coffea arabica* L.). **Ciência e Agrotecnologia**, Lavras, MG, v. 25, n. 5, p. 1248-1252, 2001.
- CINTRA, M. E.; MEIRA, C. A. A.; MONARD, M. C.; CAMARGO, H. A.; RODRIGUES, L. H. A. The use of fuzzy decision trees for coffee rust warning in Brazilian crops. In: INTERNATIONAL CONFERENCE ON INTELLIGENT SYSTEMS DESIGN AND APPLICATIONS, 11., 2011, Córdoba. **Proceedings...** Piscataway: IEEE, 2011. p. 1347-1352. ISDA 2011.
- COAKLEY, S. M. Variation in climate and prediction of disease in plants. **Annual Review of Phytopathology**, Palo Alto, v. 26, p. 163-181, Sept. 1988. DOI:10.1146/annurev.py.26.090188.001115.
- CONTRERAS-MEDINA, L. M.; OSORNIO-RIOS, R. A.; TORRES-PACHECO, I.; ROMERO-TRONCOSO, R. J.; GUEVARA-GONZÁLEZ, R. G.; MILLAN-ALMARAZ, J. R. Smart sensor for real-time quantification of common symptoms present in unhealthy plants. **Sensors**, Basel, v. 12, n. 1, p. 784-805, Jan. 2012. DOI. 10.3390/s120100784.
- COSTA, R. V.; ZAMBOLIM, L.; VALE, F. X. R.; MIZUBUTI, E. S. G. Previsão da requeima da batateira. **Fitopatologia Brasileira**, Brasília, DF, v. 27, n. 4, p. 349-354, 2002.
- DEL PONTE, E. M.; GODOY, C. V.; LI, X.; YANG, X. B. Predicting severity of Asian soybean rust epidemics with empirical rainfall models. **Phytopathology**, St.Paul, v. 96, n. 7, p. 797-803, July, 2006. DOI: 10.1094/PHYTO-96-0797.
- DE WOLF, E. D.; MADDEN, L. V.; LIPPS, P. E. Risk assessment models for wheat Fusarium head blight epidemics based on within-season weather data. **Phytopathology**, St.Paul, v. 93, n. 4, p. 428-435, Apr. 2003. DOI: 10.1094/PHYTO.2003.93.4.428.
- DI GIROLAMO NETO, C. D.; MEIRA, C. A. A.; RODRIGUES, L. H. A. Avaliação de modelos de alerta da ferrugem do cafeeiro para lavouras com alta carga pendente de frutos. In: CONGRESSO BRASILEIRO DE PESQUISAS CAFEEIRAS, 38., 2012, Caxambu. **Anais...** Brasília, DF: Embrapa Café, 2012. p. 120-121.
- DI GIROLAMO NETO, C. D.; RODRIGUES, L. H. A.; MEIRA, C. A. A. Modelos de predição da ferrugem do cafeeiro (*Hemileia vastatrix* Berkeley & Broome) por técnicas de mineração de dados. **Coffee Science**, Lavras, MG, v. 9, n. 3, p. 408-418, jul./set. 2014.
- FERNANDES, J. M.; PAVAN, W.; SANHUEZA, R. M. V. Sisalert a generic web-based plant disease forecasting system. In: INTERNATIONAL CONFERENCE ON INFORMATION AND COMMUNICATION TECHNOLOGIES IN AGRICULTURE, FOOD AND ENVIRONMENT, 5., 2011, Skiathos Island. **Proceedings...** Skiathos: HAICTA, 2011. v. 1. p. 225-233.
- FIALHO, F. B.; GARRIDO, L. da R.; BOTTON, M.; MELO, G. W. B. de; FAJARDO, T. V. M.; NAVES, R. de L. **Diagnóstico de doenças e pragas na cultura da videira usando o sistema especialista Uzum**. Bento Gonçalves: Embrapa Uva e Vinho, 2012. 4 p. (Embrapa Uva e Vinho. Comunicado técnico, 128). Disponível em: <<http://ainfo.cnptia.embrapa.br/digital/bitstream/item/73877/1/cot128.pdf>>. Acesso em: 12 out. 2014.
- FOSTER, M.; ITO, R. P. M.; WARD, R. Detection and counting of uneaten food pellets in a sea cage using image-analysis. **Aquacultural Engineering**, London, v. 14, n. 3, p. 251-269, 1995. DOI: 10.1016/0144-8609(94)00006-M.
- GARÇON, C. L. P., ZAMBOLIM, L., MIZUBUTI, E. S. G., VALE, F. X. R.; COSTA, H. Controle da ferrugem do cafeeiro com base no valor de severidade. **Fitopatologia Brasileira**, Brasília, DF, v. 29, n. 5, p. 486-491, 2004.
- GENT, D. H.; SCHWARTZ, H. F. Validation of potato early blight disease forecast models for Colorado using various sources of meteorological data. **Plant Disease**, St Paul, v. 87, n. 1, p. 78-84, Jan. 2003. DOI: 10.1094/PDIS.2003.87.1.78.

- HARDWICK, N. V. Disease forecasting. In: COOKE, B. M.; JONES, D. G.; KAYE B. (Ed.) **The epidemiology of plant diseases**. 2nd ed. Dordrecht: Springer, 2006. p. 239-267.
- HAU, B.; KRANZ, J. Mathematics and statistics for analysis in epidemiology. In: KRANZ, J. (Ed.) **Epidemics of plant diseases: mathematical analysis and modeling**. 2nd ed. Berlin: Springer-Verlag, 1990. p. 12-52.
- HUANG, K. Y. Application of artificial neural network for detecting Phalaenopsis seedling diseases using color and texture features. **Computers and Electronics in Agriculture**, New York, v. 57, n. 1, p. 3-11, May 2007. DOI: 10.1016/j.compag.2007.01.015.
- JENSEN, R. E.; BOYLE, L. W. A technique for forecasting leafspot on peanuts. **Plant Disease Reporter**, Beltsville, v. 50, n. 11, p. 810-814, 1966.
- KUSHALAPPA, A. C. Biology and epidemiology. In: KUSHALAPPA, A. C.; ESKES, A. B. (Ed.) **Coffee rust: epidemiology, resistance, and management**. Boca Raton, Florida: CRC Press, 1989a. p. 13-80.
- KUSHALAPPA, A. C. Rust management: an epidemiological approach and chemical control. In: KUSHALAPPA, A. C.; ESKES, A. B. (Ed.) **Coffee rust: epidemiology, resistance, and management**. Boca Raton, Florida: CRC Press, 1989b. p. 81-139.
- KUSHALAPPA, A. C. Epidemiologia da ferrugem do cafeeiro sob alta densidade de plantio: um enfoque de sistema. In: SIMPÓSIO INTERNACIONAL SOBRE CAFÉ ADENSADO, 1994, Londrina. **Anais...** Londrina: IAPAR, p. 131-147, 1994.
- KUSHALAPPA, A. C.; ESKES, A. B. Advances in coffee rust research. **Annual Review of Phytopathology**, Palo Alto, v. 27, p. 503-531, Sept. 1989.
- KUSHALAPPA, A. C.; AKUTSU, M.; LUDWIG, A. Application of survival ratio for monocyclic process of *Hemileia vastatrix* in predicting coffee rust infection rates. **Phytopathology**, St. Paul, v. 73, n. 1, p. 96-103, 1983. DOI: 10.1094/Phyto-73-96
- KUSHALAPPA, A. C.; HERNANDEZ, T. A.; LEMOS, H. G. Evaluation of simple and complex coffee rust forecasts to time fungicide application. **Fitopatologia Brasileira**, Brasília, DF, v. 11, p. 515-26, out. 1986.
- KUSHALAPPA, A. C.; AKUTSU, M.; OSEGUERA, S. H.; CHAVES, G. M.; MELLES, C. Equations for predicting the rate of coffee rust development based on net survival ratio for monocyclic process of *Hemileia vastatrix*. **Fitopatologia Brasileira**, Brasília, DF, v. 9, p. 255-271, jun. 1984.
- LLORET, J.; BOSCH, I.; SENDRA, S.; SERRANO, A. A wireless sensor network for vineyard monitoring that uses image processing. **Sensors**, Basel, v. 11, n. 6, p. 6165-6196, 2011. DOI: 10.3390/s110606165.
- LUACES, O.; RODRIGUES, L. H. A.; MEIRA, C. A. A.; BAHAMONDE, A. Using nondeterministic learners to alert on coffee rust disease. **Expert Systems with Applications**, New York, v. 38, n. 11, p. 14276-14283, Oct. 2011. DOI: 10.1016/j.eswa.2011.05.003.
- LUACES, O.; RODRIGUES, L. H. A.; MEIRA, C. A. A.; QUEVEDO, J. R.; BAHAMONDE, A. Viability of an alarm predictor for coffee rust disease using interval regression. In: INTERNATIONAL CONFERENCE ON INDUSTRIAL ENGINEERING AND OTHER APPLICATIONS OF APPLIED INTELLIGENT SYSTEMS, 23., 2010, Cordoba. **Trends in applied intelligent systems: proceedings**. Berlin: Springer, 2010. p. 337-346. (Lecture notes in artificial intelligence, 6097; Lecture notes in computer science).
- MADDEN, L. V.; ELLIS, M. A. How to develop plant disease forecasters. In: KRANZ, J.; ROTEM, J. (Ed.) **Experimental techniques in plant disease epidemiology**. Berlin: Springer-Verlag, 1988. p. 191-208.
- MADDEN, L.; PENNYPACKER, S. P.; MAC NAB, A. A. FAST, a forecast system for Alternariasolani on tomato. **Phytopathology**, St. Paul, v. 68, p. 1354-1358, 1978.
- MAGAREY, R. D.; SEEM, R. C.; RUSSO, J. M.; ZACK, J. W.; WAIGHT, K. T.; TRAVIS, J. W.; OUDEMANS, P. V. Site-specific weather information without on-site sensors. **Plant Disease**, St Paul, v. 85, n. 12, p. 1216-1226, Dec. 2001.

- MASSRUHÁ, S. M. F. S.; LIMA, H. P. Uma nova abordagem para diagnóstico de doenças via web. In: CONGRESSO BRASILEIRO DE AGROINFORMÁTICA, 8., 2011, Bento Gonçalves. **Anais...** Florianópolis: UFSC; Pelotas: UFPel, 2011. Disponível em: <http://ainfo.cnptia.embrapa.br/digital/bitstream/item/46276/1/89850-1.pdf>. Acesso em 10 out. 2014.
- MASSRUHÁ, S. M. F. S.; DUTRA, J. P.; CRUZ, S. A. B. D. **Uma abordagem orientada a objetos para ambiente de diagnóstico remoto**. Campinas: Embrapa Informática Agropecuária, 2007a. 33 p. (Embrapa Informática Agropecuária. Boletim de pesquisa e desenvolvimento, 15). Disponível em: <http://ainfo.cnptia.embrapa.br/digital/bitstream/CNPTIA/11553/1/bp15.pdf>. Acesso em: 10 out. 2014.
- MASSRUHÁ, S. M. F. S.; DUTRA, J. P.; CRUZ, S. A. B.; SANDRI, S.; WAINER, J. Uma plataforma orientada a objetos para desenvolvimento de sistemas especialistas de diagnóstico de doenças via Web. In: CONGRESSO BRASILEIRO DE AGROINFORMÁTICA, 6., 2007, São Pedro, SP. **Anais...** Campinas: Embrapa Informática Agropecuária, 2007b. p. 176-180. SBIAgro 2007.
- MASSRUHÁ, S. M. F. S.; DUTRA, J. P.; CRUZ, S. A. B.; SANDRI, S.; WAINER, J.; MORANDI, M. A. B. An object-oriented framework for virtual diagnosis. In: BIENNIAL CONFERENCE OF EUROPEAN FEDERATION OF IT IN AGRICULTURE, 6., 2007c, Glasgow. **Environmental and Rural Sustainability through ICT: proceedings**. Glasgow: Glasgow Caledonian University, 2007c. 6 p. Disponível em: <http://ainfo.cnptia.embrapa.br/digital/bitstream/item/101814/1/2007AA-092.pdf>. Acesso em: 10 out. 2014.
- 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) - Universidade Estadual de Campinas, Campinas.
- MEIRA, C. A. A.; RODRIGUES, L. H. A. Modelos em árvore de decisão para alerta da ferrugem do cafeeiro. In: SIMPÓSIO DE PESQUISA DOS CAFÉS DO BRASIL, 6., 2009, Vitória. **Inovação científica, competitividade e mudanças climáticas: anais...** Vitória: Consórcio Pesquisa Café, 2009. Não paginado.
- MEIRA, C. A. A.; RODRIGUES, L. H. A.; MORAES, S. A. Análise da epidemia da ferrugem do cafeeiro com árvore de decisão. **Tropical Plant Pathology**, Brasília, DF, v. 33, n. 2, p. 114-124, Mar./Apr. 2008.
- MEIRA, C. A. A.; RODRIGUES, L. H. A.; MORAES, S. A. Modelos de alerta para o controle da ferrugem-do-cafeeiro em lavouras com alta carga pendente. **Pesquisa Agropecuária Brasileira**, Brasília, DF, v. 44, p.233-242, mar. 2009. Disponível em: http://ainfo.cnptia.embrapa.br/digital/bitstream/item/93878/1/66.pdf. Acesso em: 10 out. 2014.
- MICHALSKI, R. S.; DAVIS, J. H.; BISHT, V. S.; SINCLAIR, J. B. PLANT/DS: an expert consulting system for the diagnosis of soybean diseases. In: EUROPEAN CONFERENCE ON ARTIFICIAL INTELLIGENCE, 1982, Orsay. **Proceedings...** Kaiserslautern: Universität Kaiserslautern, 1982. p. 82-84.
- MOLINEROS, J. E.; DE WOLF, E. D.; FRANCL, L.; MADDEN, L.; LIPPS, P. Modeling epidemics of fusarium head blight: trials and tribulations. **Phytopathology**, St.Paul, v. 95, p. 71, 2005.
- MORAES, S. A. **A ferrugem do cafeeiro: importância, condições predisponentes, evolução e situação no Brasil**. Campinas: Instituto Agrônomo, 1983. 50 p. (IAC. Circular, 119).
- MORAES, S. A.; SUGIMORI, M. H.; RIBEIRO, I. J. A.; ORTOLANI, A. A.; PEDRO JR., M. J. Período de incubação de *Hemileia vastatrix* Berk. et Br. em três regiões do Estado de São Paulo. **Summa Phytopathologica**, Piracicaba, v. 2, n. 1, p. 32-38, 1976.
- PARVIN JUNIOR., D. W.; SMITH, D. H.; CROSBY, F. L. Development and evaluation of a computerized forecasting method for Cercospora leaf spot of peanuts. **Phytopathology**, St. Paul, v. 64, p. 385-388, Mar. 1974.
- PAUL, P. A.; MUNKVOLD, G. P. A model-based approach to preplanting risk assessment for gray leaf spot of maize. **Phytopathology**, St. Paul, v. 94, n. 12, p. 1350-1357, Dec. 2004.
- PAUL, P. A.; MUNKVOLD, G. P. Regression and artificial neural network modeling for the prediction of gray leaf spot of maize. **Phytopathology**, St. Paul, v. 95, n. 4, p. 388-396, Apr. 2005. DOI: 10.1094/PHYTO-95-0388.

PEDRO JÚNIOR, M. J.; MORAES, S. A.; GODOY, I. J. Agrometeorological forecasting method for cercospora leaf spot in peanuts. **Fitopatologia Brasileira**, Brasília, DF, v. 19, n. 1, p. 69-73, 1994.

PÉREZ-ARIZA, C. B.; NICHOLSON, A. E.; FLORES, M. J. Prediction of coffee rust disease using bayesian networks. In: EUROPEAN WORKSHOP ON PROBABILISTIC GRAPHICAL MODELS, 6., 2012, Granada. **Proceedings...** Granada University, 2012, p. 259-266.

PINTO, A. C. S.; POZZA, E. A.; SOUZA, P. E.; POZZA, A. A. A.; TALAMINI, V.; BOLDINI, J. M.; SANTOS, F. S. Descrição da epidemia da ferrugem do cafeeiro com redes neuronais. **Fitopatologia Brasileira**, Brasília, DF, v. 27, n. 5, p. 517-524, out. 2002.

POZZA, E. A.; MAFFIA, L. A.; SILVA, C. A. B.; BRAGA, J. L.; CERQUEIRA, F. G. TOMEX-UFV: um sistema especialista para diagnose de doenças do tomateiro. In: CONGRESSO DA SBIAGRO, 1. AGROSOFT 97, 1997, Belo Horizonte. **Anais...** [S.l.:s.n.], 1997.

PYDIPATI, R.; BURKS, T. F.; LEE, W. S. Identification of citrus disease using color texture features and discriminant analysis. **Computers and Electronics in Agriculture**, New York, v. 52, n. 1-2, p. 49-59, Jun. 2006. DOI: 10.1016/j.compag.2006.01.004.

QUINLAN, J. R. **C4.5**: programs for machine learning. San Francisco: Morgan Kaufmann, 1993. 302 p. il. (The Morgan Kaufmann series in machine learning).

REIS, E. M.; BRESOLIN, A. C. R. Sistemas de previsão de doenças de plantas. In: REIS, E. M. (Org.) **Previsão de doenças de plantas**. Passo Fundo: UPF, 2004. p. 155-287.

SENA JUNIOR, D. G.; PINTO, F. A. C.; QUEIROZ, D. M.; VIANA, P. A. Fallarmywormdamagedmaizeplantidentificationusing digital images. **Biosystems Engineering**, London, v. 85, n. 4, p. 449-454, Aug. 2003.

SILVA, F. C.; MASSRUHÁ, S. M. F. S.; DEUS, R. S.; SANTOS, A. D.; BARBIERI, V.; CRUZ, S. A. B.; MALAVOLTA, E. A web-based expert system for diagnosis of nutritional deficiency in sugarcane. **Bioenergia em Revista: Diálogos**, Piracicaba, v. 1, n. 1, p. 26-48, jan./jun. 2011.

SHTIENBERG, D.; ELAD, Y. Incorporation of weather forecasting in integrated, biological-chemical management of *Botrytis cinerea*. **Phytopathology**, St. Paul, v. 87, n. 3, p. 332-340, Mar. 1997.

SMITH, M.; BARTON, M.; BASS, M.; BRANSCHOFKY, M.; MCCLELLAN, G.; STUVE, D.; TANSLEY, R.; WALKER, J. H. DSpace: an open source dynamic digital repository. **D-Lib Magazine**, v. 9, n. 1, 2003. Disponível em: <<http://www.dlib.org/dlib/january03/smith/01smith.html>>. Acesso em: 14 out. 2014.

SOUZA, R. T.; FORCELINI, C. A.; REIS, E. M.; CALVETE, E. O. Validação de dois sistemas de previsão para a queima das folhas da cenoura. **Fitopatologia Brasileira**, Brasília, DF, v. 27, n. 1, p. 87-90, jan./fev. 2002.

STORY, D.; KACIRA, M.; KUBOTA, C.; AKOGLU, A.; AN, L. Lettuce calcium deficiency detection with machine vision computed plant features in controlled environments. **Computers and Electronics in Agriculture**, New York, v. 74, n. 2, p. 238-243, Nov. 2010.

THAMADA, T. T.; GIROLAMO NETO, C. D.; MEIRA, C. A. A. Sistema de alerta da ferrugem do cafeeiro. In: SIMPÓSIO DE PESQUISA DOS CAFÉS DO BRASIL, 8., 2013, Salvador. **Pesquisa cafeeira: sustentabilidade e inclusão social**: anais. Brasília, DF: Embrapa Café, 2013. Disponível em: <<http://ainfo.cnptia.embrapa.br/digital/bitstream/item/93878/1/66.pdf>>. Acesso em: 10 out. 2014.

WALLIN, J. R. Summary of recent progress in predicting late blight epidemics in United States and Canada. **American Potato Journal**, Orono, v. 39, p. 306-312, 1962.

WITTEN, I. H.; FRANK, E.; HALL, M. A. **Data mining**: practical machine learning tools and techniques. 3rd ed. San Francisco: Morgan Kaufmann, 2011. 629 p. (Morgan Kaufmann series in data management systems).

XU, G.; ZHANG, F.; SHAH, S. G.; YE, Y.; MAO, H. Use of leaf color images to identify nitrogen and potassium deficient tomatoes. **Pattern Recognition Letters**, Amsterdam, v. 32, n. 11, p. 1584-1590, Aug. 2011. DOI: 10.1016/j.patrec.2011.04.020.

XU, X.; HARRIS, D. C.; BERRIE, A. M. Modeling infection of strawberry flowers by *Botrytis cinerea* using field data. **Phytopathology**, St. Paul, v. 90, n. 12, p. 1367-1374, Dec. 2000. DOI: 10.1094/PHTO.2000.90.12.1367.

ZADOKS, J. C. A quarter century of disease warning, 1958 - 1983. **Plant Disease**, Saint Paul, v. 68, n. 4, p. 352-355, 1984. DOI: 10.1094/PD-68-352.

ZAMBOLIM, L.; VALE, F. X. R.; COSTA, H.; PEREIRA, A. A.; CHAVES, G. M. Epidemiologia e controle integrado da ferrugem-do-cafeeiro. In: ZAMBOLIM, L. (Ed.). **O estado da arte de tecnologias na produção de café**. Viçosa, MG: Suprema Gráfica, 2002. p. 369-449.

ZAMBOLIM, L.; VALE, F. X. R.; ZAMBOLIM, E. M. Doenças do cafeeiro. In: KIMATI, H.; AMORIM, L.; REZENDE, J. A. M.; BERGAMIN FILHO, A.; CAMARGO, L. E. A. (Ed.). **Manual de fitopatologia**: doenças das plantas cultivadas. 4. ed. São Paulo: Agronômica Ceres, 2005. v. 2. p. 165-180.

ZHOU, Z.; ZANG, Y.; LI, Y.; ZHANG, Y.; WANG, P.; LUO, X. Rice plant-hopper infestation detection and classification algorithms based on fractal dimension values and fuzzy C-means. **Mathematical and Computer Modelling**, Oxford, v. 58, n. 3-4, p. 701-709, Aug. 2011.