

Uso de metadados para a geração de mapas de zona de manejo de pragas em pomar: um estudo de caso

Use of metadata to generate pest management zone maps in orchards: a case study

Luciano Gebler¹, Eduardo Antonio Speranza², Célia Regina Grego³

¹ Pesquisador, Embrapa Uva e Vinho, Vacaria (RS), Brasil, luciano.gebler@embrapa.br

² Pesquisador, Embrapa Agricultura Digital, Campinas (SP), Brasil, eduardo.speranza@embrapa.br

³ Pesquisadora, Embrapa Agricultura Digital, Campinas (SP), Brasil, celia.grego@embrapa.br

RESUMO

Grande parte da fruticultura no Brasil dispõe de algum tipo de controle de rastreabilidade, necessitando de coleta e manutenção de algum tipo de dado sobre o processo produtivo. Esse controle é uma excelente oportunidade de agregar valor a um conjunto de dados que, na maioria dos casos, é armazenado ou descartado sem uma exploração mais aprofundada. Em tempos de agricultura de precisão e digital, o uso alternativo de dados disponíveis, coletados de fontes indiretas, pode contribuir para a economicidade do processo produtivo, bastando um novo olhar sobre uma valiosa fonte de informação. O controle de pragas por manejo integrado é uma prática comumente utilizada em pomares de maçãs no sul do Brasil e é uma importante ferramenta que impacta diretamente na melhoria da qualidade e no aumento da produtividade. O uso de metadados originados de processos de certificação da qualidade e rastreabilidade dos pomares permite analisar as tendências históricas da ocorrência de algumas pragas, auxiliando na tomada de decisão e no gerenciamento da produção. Neste trabalho, o objetivo foi analisar a incidência de diferentes tipos de pragas aéreas ao longo de três safras em um pomar comercial de maçãs do município de Vacaria-RS, a partir de dados coletados para outras finalidades. Com isso, foi possível gerar mapas de regiões com diferentes níveis de susceptibilidade a pragas utilizando métodos de identificação da variabilidade espacial, como a geoestatística e os algoritmos de agrupamento específicos para dados de contagem. O exemplo do processo de análise descrito neste trabalho pode servir como auxílio para análises de risco futuras, com base em dados indiretos. Também demonstra o potencial e o valor dos dados coletados e atualmente armazenados para outras finalidades, mas que podem ser reutilizados para a adoção da fruticultura de precisão, trazendo benefícios de manejo e redução de custos ao produtor.

Palavras-chave: fruticultura de precisão; rastreabilidade; banco de dados.

ABSTRACT

A large part of fruit growing in Brazil has some type of traceability control, requiring the collection and maintenance of some type of data about the production process. This control is an excellent opportunity to add value to a data set that, in most cases, is stored or discarded without further exploration. In times of precision and digital agriculture, the alternative use of available data, collected from indirect sources, can contribute to the economy of the production process, simply by taking a new look at a valuable source of information. Pest control through integrated management is a practice commonly used in apple orchards in southern Brazil and is an important tool that directly impacts quality improvement and increased productivity. The use of metadata from orchard quality and traceability certification processes enable to analyze historical trends in the occurrence of some pests, assisting in decision-making and production management. In this work, the objective was to analyze the incidence of different types of aerial pests over three harvests in a commercial apple orchard in Vacaria, RS, Brazil, based on data collected for other purposes. With this, it was possible to generate maps of regions with different levels of susceptibility to pests using methods for identifying spatial variability such as geostatistics and specific grouping algorithms for count data. The example of the analysis process described in this work can help future risk analyzes based on indirect data. It also demonstrates the potential and value of data collected and currently

<https://doi.org/10.4322/978-65-86819-38-0.1000032>

 Este é um capítulo publicado em acesso aberto (Open Access) sob a licença Creative Commons Attribution-NonCommercial-NoDerivatives, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que sem fins comerciais, sem alterações e que o trabalho original seja corretamente citado.

stored for other purposes, but they can be reused for the adoption of precision fruit growing, bringing management benefits and cost reduction to the producer.

Keywords: precision fruit crop; traceability; database.

1 INTRODUÇÃO

A exploração econômica nas cultivares de macieira expandiu significativamente nos últimos anos, principalmente nos estados de Santa Catarina e Rio Grande do Sul, sendo a produção autossuficiente para o consumo interno, além das exportações para mercados altamente exigentes. Diante desse cenário, empresas produtoras caminham em direção à produção eficiente com qualidade e sustentabilidade, minimizando os efeitos negativos à saúde e ao meio ambiente pela redução de utilização de agrotóxicos (Fioravanço; Santos, 2013).

Com o aumento de produtividade e a exigência em qualidade da fruta, houve a necessidade da adoção de um sistema que auxiliasse no gerenciamento e manejo das áreas, surgindo então a Produção Integrada de Maçãs (PIM), programa de certificação de qualidade que exigia a implantação de um sistema para auxílio na tomada de decisões relativas ao uso de agrotóxicos com base no monitoramento de pragas (Kovaleski; Ribeiro, 2002). Nesse contexto, nos estados de Santa Catarina e Rio Grande do Sul, atualmente responsáveis por cerca de 97% da produção nacional de maçãs (Lazzarotto, 2018), a adoção da PIM é uma realidade mesmo que, devido ao custo da fase de certificação por terceiros, poucos produtores finalizem o processo com o selo, apesar de a quase totalidade do setor produtivo seguir o protocolo técnico na íntegra.

No caso da PIM, esta proporciona ao produtor uma atualização dinâmica da incidência de pragas na sua área de produção, facilitando o controle das populações e obrigando ainda que os produtores armazenem seus dados de monitoramento por um período mínimo de cinco anos, caso haja necessidade de auditorias extras pelo programa. A partir dessa exigência, os produtores passaram a ter a disponibilidade de planilhas de dados geográficos e temporais em seus pomares (Fioravanço; Santos, 2013). Durante a safra, estes dados são normalmente utilizados semanalmente para manejo de pragas, mas, por falta de ferramentas ou orientação, acabam sem utilização posterior, sendo estocados tanto de maneira física (cadernos de campo) como virtual (planilhas eletrônicas).

Outros sistemas de certificação da qualidade além da PIM são utilizados pelo setor, como a GlobalGAP ou BRC (*British Retail Consortium*), que também demandam as mesmas regras de coleta e armazenamen-

to de dados para processos de auditoria, resultando em um enorme potencial para a geração de informações indiretas que podem apoiar o gerenciamento do setor produtivo por meio da interpretação dos novos resultados pela utilização de técnicas de fruticultura de precisão.

Além destes sistemas de certificação comercial, a partir da publicação da Instrução Normativa Conjunta nº 02/2018 pelo MAPA e pela ANVISA, o processo de rastreabilidade de frutas e hortaliças se tornou obrigatório no Brasil, exigindo que os produtores coletem e mantenham uma série de dados sobre suas culturas para fins de eventuais auditorias. Este banco de dados, atualmente coletado com objetivos legais, também pode servir para fins de apoio à gestão da propriedade por meio de seu uso na forma de metadados.

A incidência de pragas em uma área produtora de maçãs é verificada com a instalação de armadilhas para amostragem em locais específicos em cada safra. Com a obtenção desses dados ao longo dos anos, é possível identificar regiões mais susceptíveis e estabelecer o manejo adequado para o controle da população de pragas. Consequentemente, essas ações direcionadas podem favorecer o investimento em áreas nas quais há possibilidade de serem obtidos frutos de melhor qualidade. Entretanto, para que esses dados sejam transformados em informações acessíveis e facilmente interpretáveis por parte dos produtores e órgãos reguladores, é conveniente o uso de técnicas computacionais, como a inteligência artificial. Nesse contexto, a utilização de algoritmos de aprendizado de máquina para a predição de infestação de pragas e doenças na agricultura é algo que vem sendo bastante investigado na literatura. Na fruticultura, especialmente para maçãs, os trabalhos ainda são bastante recentes, como no caso de Boniecki et al. (2015), Brunelli et al. (2019) e Wrzesień et al. (2019), que desenvolveram métodos de monitoramento direto das pragas através da coleta de dados específicos para seus modelos.

Diferentemente das aplicações acima, mapas de distribuição espacial de pragas podem ser obtidos a partir de dados coletados em armadilhas instaladas em locais específicos da lavoura. Nesse caso, técnicas que identificam a variabilidade espacial podem ser muito úteis para garantir melhor precisão no mapeamento, como a utilização da ferramenta de análise

geoestatística. Por se tratarem de dados não rotulados previamente, a classificação desses mapas, considerando diferentes classes de incidência de pragas, pode ser realizada a partir de algoritmos de aprendizado de máquina não supervisionados, conhecidos como algoritmos de agrupamento. Esses algoritmos têm como principal objetivo agrupar amostras de dados de maneira natural, ou seja, fazendo com que pares de amostras mais similares com relação ao domínio da aplicação sejam alocadas em um mesmo grupo, e pares de amostras menos similares sejam alocadas em grupos distintos (Jain; Dubes, 1988).

Portanto, o objetivo deste trabalho foi demonstrar o potencial da reutilização de dados indiretos por meio de técnicas de geoestatística e algoritmo de agrupamento, para analisar a incidência de diferentes tipos de pragas aéreas ao longo de três safras em um pomar comercial de maçãs do município de Vacaria-RS, permitindo a construção de mapas de risco com diferentes níveis de susceptibilidade às pragas como apoio à tomada de decisão.

2 MATERIAL E MÉTODOS

Os dados utilizados neste trabalho foram obtidos em um pomar comercial de macieiras (*Malus domestica*) da empresa Agropecuária Schio LTDA, em Vacaria, Rio Grande do Sul, Brasil (50°49'13''O; 28°28'43''S), com área total de abrangência de aproximadamente 305 ha, com cultivares Gala, Fuji e seus respectivos clones. A empresa segue normas específicas de Manejo Integrado de Pragas (MIP), aplicando agrotóxicos somente em locais nos quais o nível de pragas obtido, conforme as informações coletadas das armadilhas,

é alto o suficiente para justificar tal aplicação. Foram disponibilizadas 248 armadilhas georreferenciadas com GPS RTK (Figura 1a), capazes de capturar as pragas lagarta-enroladeira (*Bonagota salubricola*) e mariposa-oriental (*Grapholita molesta*), com 94 armadilhas cada (Figura 1b), e mosca-das-frutas (*Anastrepha fraterculus*), com 60 armadilhas (Figura 1c).

As contagens de ocorrências realizadas nas armadilhas referenciadas na Figura 1a referem-se às safras de 2011/2012, 2012/2013 e 2013/2014. Candeia et al. (2016) realizaram em seu trabalho um pré-processamento de binarização desse mesmo conjunto de dados, atribuindo valor 1 para amostras com um número acima do mínimo e valor 0 para amostras com um número abaixo do mínimo de ocorrências determinadas para a utilização de agrotóxicos, segundo as normas do MIP utilizado pela empresa. Entretanto, para este trabalho, foram utilizados os valores absolutos de contagem, com o intuito de viabilizar a utilização de algoritmos de agrupamento de dados para realização das análises.

2.1 Análise Geoestatística

Inicialmente, a partir dos dados de incidência de pragas, foi realizada análise geoestatística para verificar a existência de dependência espacial por meio do cálculo e ajuste dos semivariogramas escalonados ajustados às funções de melhor correlação. O escalonamento dos semivariogramas foi utilizado conforme Vieira et al. (2010), com a finalidade de elaborar todos os semivariogramas na mesma escala de semivariância e distância para todas as variáveis que apresentam a mesma unidade de medida e que foram amostradas na mesma área de estudo. Os parâmetros de depen-

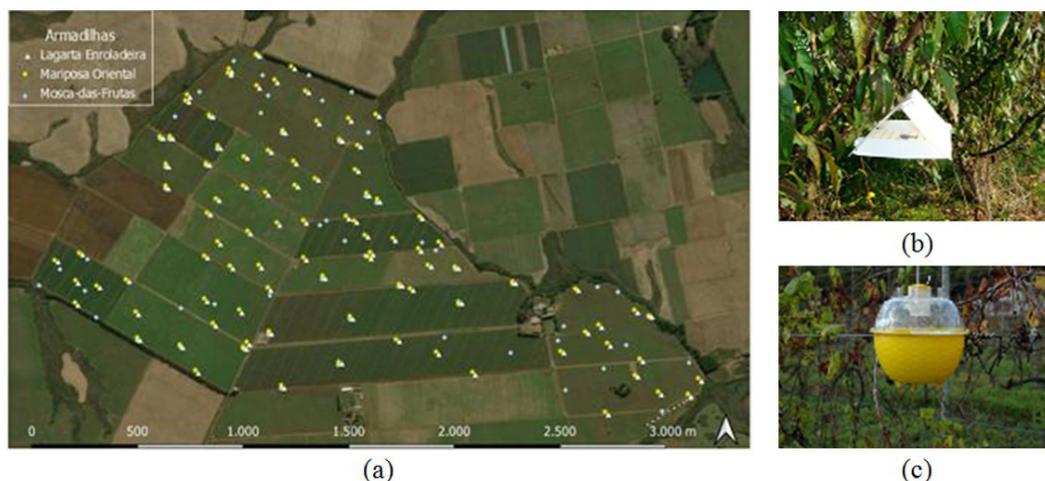


Figura 1. Área de estudo (a) e localização das armadilhas para as espécies de pragas lagarta-enroladeira, mariposa-oriental e mosca-das-frutas, e tipos de armadilhas utilizadas no campo para controle de lagarta-enroladeira e mariposa-oriental (b), além de mosca-das-frutas (c).

dência efeito pepita (Co), variância estrutural (C1) e alcance (a) foram utilizados na interpolação por krigagem. Os semivariogramas escalonados foram ajustados ao modelo esférico que melhor representou a variabilidade espacial das variáveis. A partir dos semivariogramas ajustados, os dados foram interpolados utilizando krigagem ordinária, que garantiu a interpolação sem tendência e com variância mínima, permitindo a construção de mapas de cada safra utilizando o *software* QGIS (QGIS, 2021).

2.2 Aprendizado de máquina não supervisionado

A análise geoestatística realizada permitiu identificar a presença de variabilidade espacial na área de estudo com relação à incidência de pragas, considerando os dados coletados nas três safras. A análise com uso de aprendizado de máquina permitiu a verificação da possibilidade de particionamento da área de estudo em regiões com maior e menor incidência de pragas, tomando como entrada os dados coletados nas três safras.

Devido ao fato de a aplicação deste trabalho utilizar um tipo específico de dados, assim como na análise geoestatística, foram utilizados algoritmos de agrupamento específicos para dados de contagem (não contínuos). Para identificação de qual seria o algoritmo ideal para gerar mapas de incidência para cada uma das pragas, foi utilizada a função *optCluster*, pertencente ao pacote de mesmo nome, disponível para o *software* R (Sekula et al., 2017). Essa função permite verificar, a partir dos dados e técnicas estatísticas que identificam índices de estabilidade, conectividade e homogeneidade, qual o algoritmo ideal para agrupar as amostras, considerando um total de seis algoritmos específicos para dados de contagem, bem como a quantidade ideal de grupos para classificar o conjunto de dados. Os algoritmos do reconhecimento determinístico (Ueda; Nakano, 1998) e do reconhecimento simulado (Van Laarhoven; Aarts, 1987), ambos utilizando modelo negativo binomial para contagem de dados e disponíveis no pacote *MBCluster*. *Seq* do *software* R (Si et al., 2014), foram os que apresentaram maior aderência aos conjuntos de dados utilizados. Esses algoritmos são variantes do algoritmo da maximização da expectativa (EM) (Moon, 1996) e, por se basearem em modelos tradicionais de distribuições probabilísticas para dados de contagem, se tornam adequados para simular agrupamentos para conjuntos de dados de incidência de pragas. Entretanto, como esses algoritmos são inicializados a partir de centroides aleatórios, em algum momento pode ser que o resultado convirja para um mínimo local e não global. Desse modo, cada algoritmo escolhido para cada conjunto de dados foi executado pelo menos dez vezes, em que o resultado escolhido foi aquele que apresentou a menor soma dos erros quadráticos, ou seja, a menor soma das distâncias de cada amostra com relação ao centro do grupo ao qual ela foi associada. Detalhes mais aprofundados sobre a utilização dessas técnicas nesse tipo de trabalho podem ser encontrados em Speranza et al. (2022).

Com o intuito de realizar uma comparação dos mapas obtidos por safra a partir de técnicas de geoestatística com o mapa final de incidência de pragas obtido a partir de análise de agrupamento, foi utilizada a estatística do coeficiente de *Kappa* (Mchugh, 2012). Esse coeficiente é utilizado em análise de agrupamento como um critério de validação externa, ou seja, para avaliar a concordância entre rótulos atribuídos para duas soluções distintas. O coeficiente de *Kappa* retorna valores entre 0 e 1, em que valores próximos de 1 indicam alta concordância entre os agrupamentos comparados, e valores próximos de 0 indicam baixa concordância.

Com o intuito de realizar uma comparação dos mapas obtidos por safra a partir de técnicas de geoestatística com o mapa final de incidência de pragas obtido a partir de análise de agrupamento, foi utilizada a estatística do coeficiente de *Kappa* (Mchugh, 2012). Esse coeficiente é utilizado em análise de agrupamento como um critério de validação externa, ou seja, para avaliar a concordância entre rótulos atribuídos para duas soluções distintas. O coeficiente de *Kappa* retorna valores entre 0 e 1, em que valores próximos de 1 indicam alta concordância entre os agrupamentos comparados, e valores próximos de 0 indicam baixa concordância.

3 RESULTADOS E DISCUSSÃO

Os valores mínimos, máximos e medianos de incidência das três pragas em cada uma das três safras, bem como seus coeficientes de variação e desvios padrão, podem ser visualizados na Tabela 1. Esses dados mostram dificuldades no controle da praga lagarta-

Tabela 1. Valores mínimos (mín), máximos (máx), medianos (med), desvio padrão (DP) e coeficiente de variação (CV) (%) de incidência das três pragas em macieira em três safras.

Praga / safra	2011/2012					2012/2013					2013/2014				
	mín	máx	med	DP	CV	mín	máx	med	DP	CV	mín	máx	med	DP	CV
Lagarta-enroladeira	0	2	0	0,56	132	0	12	0	2,75	186	0	9	1	2,12	134
Mariposa-oriental	0	2	0	0,34	451	0	6	0	1,05	329	0	1	0	0,02	682
Mosca-das-frutas	2	12	7	141,50	564	5	19	9,5	154,02	513	1	23	6	167,42	592

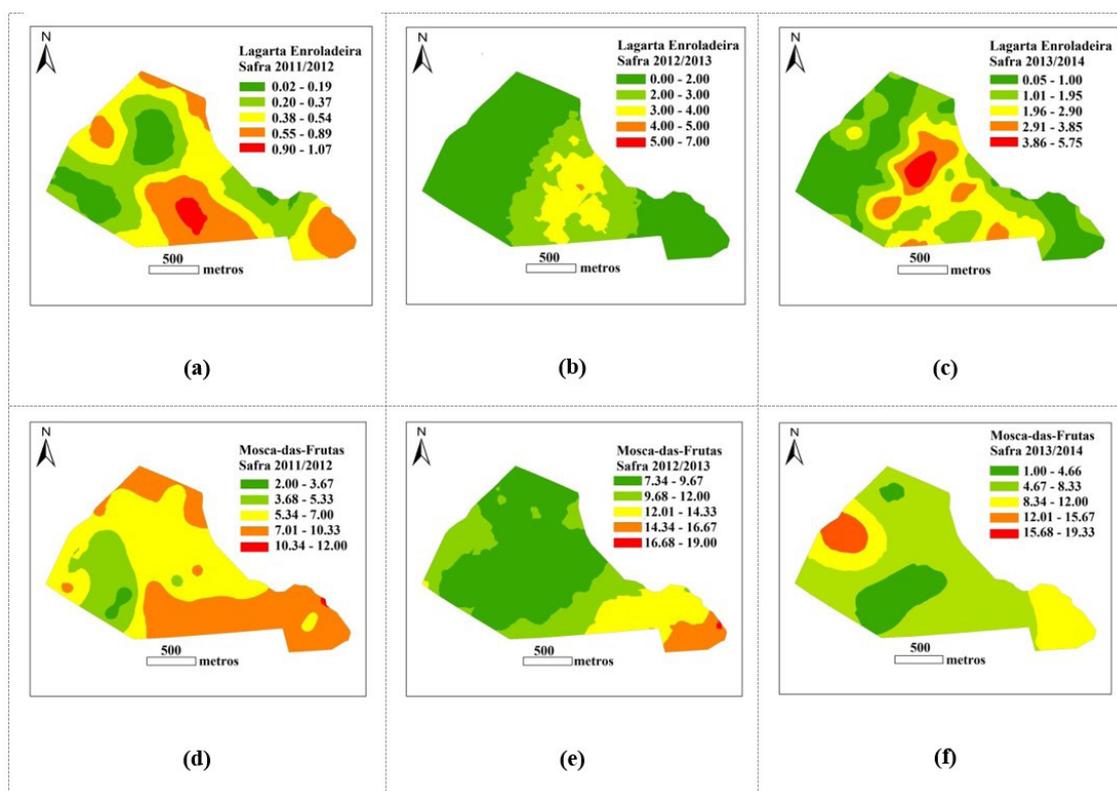


Figura 2. Mapas de evolução de ocorrências das pragas obtidos a partir de interpolação por krigagem para lagarta-enroladeira: (a) safra 2011/2012, (b) safra 2012/2013 e (c) safra 2013/2014, e para mosca-das-frutas: (d) safra 2011/2012, (e) safra 2012/2013 e (f) safra 2013/2014.

-enroladeira e, principalmente, da praga mosca-das-frutas, com aumento da incidência ao longo dos anos. A praga mariposa-oriental apresentou um pequeno aumento de incidência na safra 2012/2013, o que praticamente não influenciou no resultado final deste experimento.

Foram gerados semivariogramas escalonados para as três pragas considerando as três safras estudadas, sendo detectada dependência espacial somente para a lagarta-enroladeira e para a mosca-das-frutas com ajustes esféricos, eliminando o estudo da mariposa-oriental dos passos seguintes, nesta análise. Os mapas elaborados a partir de valores interpolados por krigagem para as variáveis que apresentaram dependência espacial encontram-se na Figura 2.

Para a praga lagarta-enroladeira, é possível perceber uma maior incidência na região central da área experimental, considerando as safras mapeadas (Figuras 2a, 2b e 2c). Nos mapas para a evolução da mosca-das-frutas, houve predominância para maiores incidências localizadas na parte inferior direita da área (Figuras 2d, 2e e 2f). Isso indica que essas regiões devem ser manejadas ou investigadas diferentemente das demais regiões da área em relação à ocorrência dessas pragas. A geração do mapa parte do princípio de que a distribuição da praga na lavoura não ocor-

re ao acaso, mas segue uma distribuição dependente do espaço e da época de cada avaliação realizada em campo, o que direciona a tomada de decisão para o controle da praga no campo.

No que diz respeito ao agrupamento de dados para obtenção de mapas com diferentes regiões de incidência de pragas, os dados de incidência, nas três safras, das pragas lagarta-enroladeira e mosca-das-frutas foram agrupados utilizando, respectivamente, os algoritmos do recozimento determinístico e do recozimento simulado, ambos com base na distribuição negativa binomial para a contagem de dados. Em ambos os casos, os melhores agrupamentos foram obtidos considerando a subdivisão em dois grupos. Devido à baixa incidência de ocorrência presente nos dados nas três safras, não foi possível encontrar algoritmo ideal para separar as incidências da praga mariposa-oriental em grupos e, desse modo, o conjunto de dados específico dessa praga foi descartado, assim como já havia ocorrido na análise geoestatística.

A partir de análises de coesão intragrupos e separação intergrupos, e considerando a existência de pontos não coincidentes de localização das armadilhas instaladas para as duas pragas, foi realizada uma fusão simples entre os agrupamentos, unindo as amostras do grupo de menor incidência para as duas

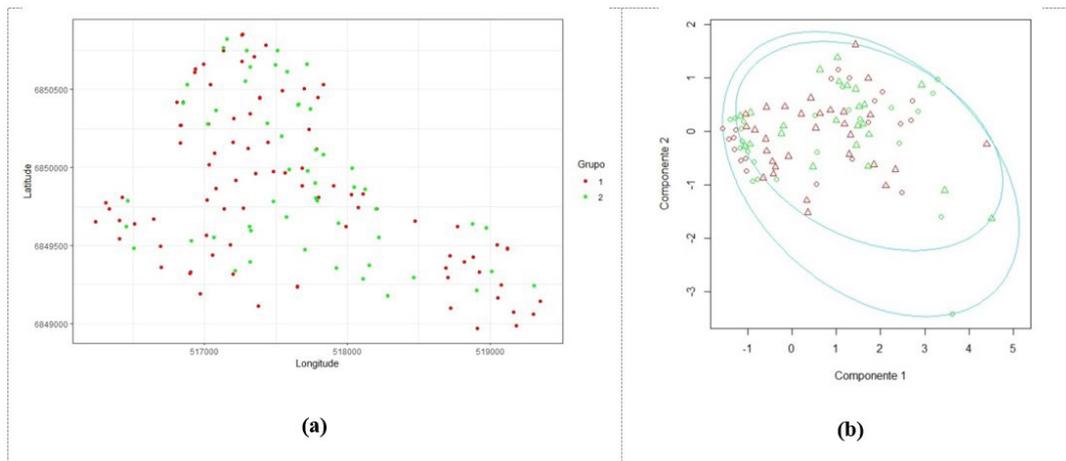


Figura 3. Agrupamento resultante da fusão dos agrupamentos gerados pelo *clusterplot* do software R: (a) agrupamento (2 grupos) exibidos em forma de mapa; (b) visualização dos grupos no espaço bidimensional (amostras nos diferentes grupos representadas por triângulos ou círculos).

pragas em um único grupo (valores medianos 1 e 8), sendo as amostras com maior incidência (valores medianos 4 e 12,5) unidas em outro grupo (Figura 3a). Entretanto, esses novos grupos, apesar de possuírem valores medianos bem distintos (1 e 5, respectivamente), não aparentam estar bem separados de uma maneira geral, pois existem muitos pontos de interseção entre eles, quando visualizados a partir das duas componentes principais (Figura 3b). Apesar de este resultado indicar certa dificuldade em separar a área de estudo em duas regiões únicas com maior e menor incidência geral de pragas, é possível observar uma área contínua de incidência acima da média para ambas as pragas, na região central do mapa (Figura 3a), cujas amostras estão representadas pelos triângulos na única região de não interseção entre os grupos (Figura 3b). Adicionalmente, na parte norte do mapa, também é possível identificar uma região contínua de incidência de pragas, influenciada principalmente pela praga mosca-das-frutas.

A partir dos dados agrupados, conforme a Figura 3a, foram utilizadas ferramentas para transformação de amostras pontuais em polígonos, disponíveis no *software* QGIS. Com isso, foi possível gerar um mapa geral contínuo de distribuição espacial das pragas da área de estudo, considerando a análise realizada a partir das três safras (Figura 4).

A partir da Figura 4, as regiões em vermelho (Grupo 1) são identificadas como de alta incidência de pragas, com valores medianos de cinco incidências por safra, e as regiões em verde (Grupo 2) são identificadas como de baixa incidência de pragas, com valores medianos de uma incidência por safra. Assim, da mesma forma que nos gráficos bidimensionais da Figura 3, é possível identificar uma região central mais contínua e ou-

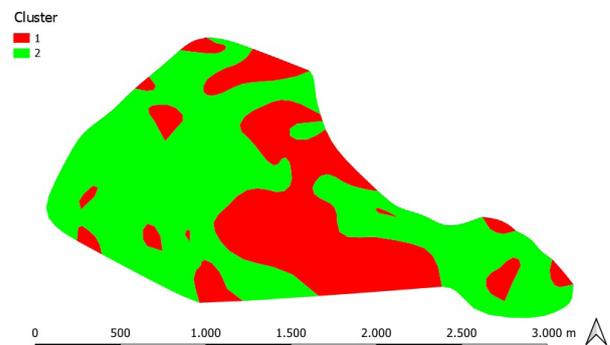


Figura 4. Mapa sugerido para gestão diferenciada de população de pragas, considerando a incidência observada em três safras das pragas lagarta-enroladeira e mosca-das-frutas. Áreas com alta incidência de pragas estão representadas em vermelho (Grupo 1) e áreas com baixa incidência de pragas estão representadas em verde (Grupo 2).

Tabela 2. Coeficiente de correlação Kappa para comparação entre mapa de incidência de pragas único, obtido a partir de análise de agrupamento, com mapas individuais obtidos a partir de geoestatística para cada praga em cada safra.

Praga / Safra	2011/2012	2012/2013	2013/2014
Lagarta-enroladeira	0,23	0,48	0,26
Mosca-das-frutas	0,36	0,47	0,35

tra região menor, localizada ao norte do mapa, com alta incidência de pragas.

Com o objetivo de verificar a concordância do mapa obtido a partir de análise de agrupamento (Figura 4) com os mapas obtidos a partir de geoestatística (Figura 2), foi utilizado o coeficiente de correlação *Kappa*. Para tanto, as cinco classes de cada um dos seis mapas da Figura 2 foram convertidas em apenas

duas classes (com maior e menor incidência de pragas), considerando a divisão de valores a partir de algoritmo de quebras naturais (Jenks, 1967). A Tabela 2 exibe os resultados de correlação obtidos com o coeficiente de correlação *Kappa*.

Seguindo a classificação utilizada para o critério de correlação *Kappa* (Viera; Garrett, 2005), para ambas as pragas, a correlação com o mapa geral obtido a partir de análise de agrupamento pode ser considerada razoável, para as safras de 2011/2012 e 2013/2014, e moderada, para a safra 2012/2013. O fato de não existirem correlações com valores de *Kappa* mais próximos de 1, que seriam classificadas como substanciais ou quase perfeitas, mostra que o algoritmo de agrupamento utilizado conseguiu fornecer uma solução mais geral que considera, de maneira justa, a dinâmica temporal de infestação de pragas, o que pode ser observado nos mapas individuais obtidos a partir da geoestatística. Adicionalmente, o não aparecimento de valores abaixo de 0,20 para o índice *Kappa* também indica que nenhum dos mapas individuais mostrou pouca correlação com a solução unificada.

A partir de novas análises georreferenciadas de fatores relacionados ao solo, à cultura e ao microclima, será possível investigar melhor as causas da incidência maior de pragas nos locais exibidos no mapa da Figura 4, priorizando um manejo mais específico como forma de reduzir os riscos de aumento da população de pragas nas safras seguintes. Adicionalmente, este conhecimento referente à variabilidade temporal do comportamento das pragas analisadas permite, ao gestor da propriedade ou ao tomador de decisão, definir melhor as estratégias de controle de pragas, buscando o manejo dos pomares com base no risco histórico das áreas, principalmente em situações nas quais há limitações de recursos ou mão de obra. Assim, é possível orientar maiores esforços para as áreas de maior risco (zonas vermelhas) em relação àquelas de menor risco (zonas verdes), e, à medida que os dados forem se acumulando, obter mais clareza a respeito das causas das variabilidades espacial e temporal do comportamento dos insetos naquela área.

A abordagem utilizada neste trabalho – apesar de os mapas contínuos obtidos a partir de interpolação de dados por krigagem não terem sido utilizados como informações para os algoritmos de agrupamento – é muito utilizada no delineamento de zonas de manejo. Portanto, essa metodologia pode ser explorada para reduzir o aparecimento das pequenas áreas com alta incidência de pragas observadas no mapa sugerido para gestão diferenciada de população de pragas

4 CONCLUSÕES

O exemplo apresentado demonstrou que o uso de dados indiretos armazenados pelo produtor ainda podem gerar informações valiosas para a sua tomada de decisão. Uma vez que a coleta e o armazenamento de dados de incidência de pragas são obrigatórios para os programas de certificação, sua aplicação na forma de metadados pode servir como base a um processo de análise de risco, trazendo benefícios de manejo e redução de custos ao produtor. A análise geoestatística permitiu verificar a presença de infestações de determinadas pragas na área de estudo, sendo da praga lagarta-enroladeira na região central e da praga mosca-das-frutas na região inferior direita. Considerando a análise de agrupamento, foi possível identificar regiões com incidência acima da média para ambas as pragas, ao mesmo tempo, principalmente na região central do mapa.

REFERÊNCIAS

- Boniecki, P.; Koszela, K.; Piekarskaboniecka, H.; Weres, J.; Zaborowicz, M.; Kujawa, S.; Majewski, A.; Raba, B. Neural identification of selected apple pests. **Computers and Electronics in Agriculture**, v. 110, p. 9-16, 2015. DOI: <http://doi.org/10.1016/j.compag.2014.09.013>.
- Brunelli, D.; Albanese, A.; d'Acunzio, D.; Nardello, M. Energy neutral machine learning based iot device for pest detection in precision agriculture. **IEEE Internet of Things Magazine**, v. 2, n. 4, p. 10-13, 2019. DOI: <http://doi.org/10.1109/IOTM.0001.1900037>.
- Candeia, M. G.; Gebler, L.; Bueno, M. B. Uso da AP no gerenciamento de metadados de longo prazo de monitoramento de pragas da macieira. In: CONGRESSO BRASILEIRO DE AGRICULTURA DE PRECISÃO (CONBAP), 2016, Goiânia. **Anais [...]**. Goiânia: AsBraAP, 2016.
- Fioravanço, J. C.; Santos, R. S. S. **Maçã: o produtor pergunta, a Embrapa responde**. Brasília, DF: Embrapa, 2013. 239 p. (Coleção 500 perguntas, 500 respostas).
- Jain, A. K.; Dubes, R. C. **Algorithms for clustering data**. Londres: Prentice-Hall, Inc., 1988.
- Jenks, G. F. The data model concept in statistical mapping. **International Yearbook of Cartography**, v. 7, p. 186-190, 1967.
- Kovaleski, A.; Ribeiro, L. G. **Manejo de pragas na produção integrada da maçã**. Bento Gonçalves: Embrapa Uva e Vinho, 2002. p. 8. (Embrapa Uva e Vinho. Circular Técnica, 34).
- Lazzarotto, J. J. **Indicadores econômicos e financeiros em sistemas típicos de produção de maçã no Brasil**. Bento Gonçalves: Embrapa Uva e Vinho, 2018.
- Mchugh, M.L. Interrater reliability: the kappa statistic. **Biochemia Medica**, v. 22, n. 3, p. 276-282, 2012. DOI: <http://doi.org/10.11613/BM.2012.031>.

- Moon, T. K. The expectation-maximization algorithm. **IEEE Signal Processing Magazine**, v. 13, n. 6, p. 47-60, 1996. DOI: <http://doi.org/10.1109/79.543975>.
- QGIS. 2021. Disponível em: <https://qgis.org>. Acesso em: 26 jul. 2021.
- Sekula, M.; Datta, S.; Datta, S. optCluster: an R package for determining the optimal clustering algorithm. **Bioinformatics**, v. 13, n. 3, p. 101-103, 2017. DOI: <http://doi.org/10.6026/97320630013101>.
- Si, Y.; Liu, P.; Li, P.; Brutnell, T. P. Model-based clustering for RNA-seq data. **Bioinformatics (Oxford, England)**, v. 30, n. 2, p. 197-205, 2014. DOI: <http://doi.org/10.1093/bioinformatics/btt632>.
- Speranza, E. A.; Grego, C. R.; Gebler, L. Analysis of pest incidence on apple trees validated by unsupervised machine learning algorithms. **Engenharia na Agricultura**, v. 30, p. 63-74, 2022. DOI: <http://doi.org/10.13083/re-veng.v30i1.12919>.
- Ueda, N.; Nakano, R. Deterministic annealing EM algorithm. **Neural Networks**, v. 11, n. 2, p. 271-282, 1998. DOI: [http://doi.org/10.1016/S0893-6080\(97\)00133-0](http://doi.org/10.1016/S0893-6080(97)00133-0).
- Van Laarhoven, P. J. M.; Aarts, E. H. L. **Simulated annealing: theory and applications**. Dordrecht: Springer, 1987. DOI: <http://doi.org/10.1007/978-94-015-7744-1>.
- Vieira, S. R.; Garcia, M. A. G.; González, A. P.; Siqueira, G. M. Variabilidade espacial e temporal do teor de água do solo sob duas formas de uso. **Bragantia**, v. 69, n. 1, p. 181-190, 2010. DOI: <http://doi.org/10.1590/S0006-87052010000100023>.
- Viera, A. J.; Garrett, J. M. Understanding interobserver agreement: the kappa statistic. **Family Medicine**, v. 37, n. 5, p. 360-363, 2005.
- Wrzesień, M.; Treder, W.; Klamkowski, K.; Rudnicki, W. R. Prediction of the apple scab using machine learning and simple weather stations. **Computers and Electronics in Agriculture**, v. 161, p. 252-259, 2019. DOI: <http://doi.org/10.1016/j.compag.2018.09.026>.