

Prediction system for appearance of boll weevil population in cotton crops

Sistema de predição de aparecimento de população de bicudo na lavoura de algodão

Marcelo G. Narciso¹, Joécio Carvalho², José Ednilson Miranda¹

¹Embrapa Arroz e Feijão (CNPAP)
Rodovia GO-462 - CEP 75375-000 – Santo Antônio de Goiás – GO – Brazil

²Livefarm Tecnologia Agropecuária
SIG Quadra 6, 2210 - Sala 101 CEP: 70610-460 – Brasília– DF– Brazil

{marcelo.narciso,jose-ednilson.miranda}@embrapa.br, joelcio.livefarm@gmail.com

Abstract. *The boll Weevil is one of the major pests of cotton in Brazil. It is a pest that is difficult to eradicate and can cause great damage to the cotton producer. Management to control this pest is essential to avoid damage. A system that predicts the appearance of boll weevils, according to daily climatic variables, is very useful to alert the producer to the appearance of weevil populations in the field, so that he can make a decision about the appropriate management to combat the boll weevil. This work shows an alert system to predict the appearance of boll weevil in cotton crops.*

Keywords: cotton, boll weevil, information system, prediction, climatic data

Resumo. *O bicudo é uma das grandes pragas do algodoeiro no Brasil. É uma praga de difícil erradicação e pode trazer grandes prejuízos ao produtor de algodão. O manejo para controle desta praga é fundamental para evitar prejuízos. Um sistema que prevê o aparecimento do bicudo, conforme variáveis climáticas diárias, é muito útil para alertar ao produtor quanto ao aparecimento da população de bicudo na lavoura, para que este possa tomar decisão sobre o manejo adequado para combater o bicudo. Este trabalho mostra um sistema de alerta para predizer o aparecimento de bicudo na lavoura de algodão.*

Palavras-chave: algodão, bicudo, sistema de informação, predição, dados climáticos.

1. Introdução

O bicudo é uma das principais pragas que atingem a lavoura de algodão. Este inseto tem coloração cinzenta ou castanha e mandíbulas afiadas, utilizadas para perfurar o botão floral e a maçã dos algodoeiros. O adulto do bicudo tem comprimento médio de sete

milímetros, com variação de três a nove milímetros, e uma largura equivalente a um terço do comprimento [CANAL_RURAL 2023]. A Figura 1 ilustra o bicudo do algodoeiro



Figura 1. Bicudo do algodoeiro. Fonte: (AGRONOVAS, 2023)

Para se ter uma idéia do quão severa é esta praga, o bicudo, no passado, atingiu áreas produtoras do Nordeste, aniquilando a produção pela dificuldade de controle químico. Mais tarde, a maior parte da produção de algodão passou para os Estados do Paraná e São Paulo. Posteriormente, devido ao ataque de bicudos, a maior parte da produção de algodão migrou para o Centro-Oeste do país [CANAL_RURAL, 2023].

Um levantamento recente, realizado com cerca de 60 produtores de algodão do Cerrado, verificou que o bicudo é a praga que exige o maior número de pulverizações, variando entre 18 e 22 aplicações por safra, nas últimas sete safras [NEGÓCIO_RURAL, 2023].

Para o combate desta praga, existem uma série de recomendações. Algumas destas recomendações seriam: destruição de raízes, caules, botões florais, flores, maçãs, carimãs e capulhos não colhidos, através do arranque e destruição dos restos culturais após 30 dias da colheita do algodão; preparo antecipado do solo em aproximadamente 40 dias para desalojar os adultos remanescentes da safra anterior; aplicação simultânea de desfolhantes e inseticidas para reduzir danos e a densidade populacional da praga; instalação de armadilhas com feromônio, 20 dias antes da semeadura e até 60 dias após a emergência das plantas, para registrar a presença de adultos na área cultivada [NEGÓCIO_RURAL, 2023].

Uma forma que ajudaria a combater esta praga seria o uso, por parte do produtor, de um sistema de alerta que consiga prever o aparecimento da população de bicudo na lavoura de algodão a qualquer momento, sem precisar de uma verificação no campo com pragueiro. Visto que variáveis climáticas tem influência quanto ao aparecimento de população do bicudo na lavoura, como pode ser visto em [Gabriel 2016], [Degrande et al 1983] e [Volpe 1993], o sistema de alerta pode ser feito com variáveis climáticas como entrada e a saída, o grau de severidade de aparecimento da praga.

Este trabalho contempla um sistema de alerta que prediz o aparecimento de população de bicudo em função de variáveis climáticas, as quais o produtor poderá obter a partir de dados de instituições como INMET (ver <https://portal.inmet.gov.br>), Agritempo (ver <https://www.agritempo.gov.br/agritempo/index.jsp>), NASA Power (ver <https://power.larc.nasa.gov/>) ou ainda, se o produtor tiver uma estação meteorológica em sua propriedade. Este sistema visa a auxiliar o produtor quanto à tomada de decisão para combate do bicudo em lavouras de algodão.

2. Material e Métodos

Para o desenvolvimento de um sistema de alerta que possa prever o aparecimento de bicudo na lavoura, é necessário correlacionar o aparecimento de bicudo, na lavoura de algodão, com variáveis climáticas. Para isso, foram feitos seis experimentos com armadilhas para bicudo em fazendas da Bahia, na zona rural dos municípios de São Desidério, Luís Eduardo Magalhães e Barreiras. A Figura 2 ilustra a disposição de armadilhas em um dos experimentos no Centro de Pesquisa e Tecnologia do Oeste da Bahia – CPTO, próximo ao município de Luís Eduardo Magalhães.



Figura 2. Disposição de armadilhas em um dos experimentos. Fonte: [LIVEFARM 2023]

Em cada plantação de algodão, foram instaladas armadilhas para captura de bicudos. Em cada semana, eram contados quantos bicudos tinham em cada armadilha e logo a seguir estes bicudos eram descartados, isto é, as armadilhas ficavam vazias. Cada armadilha tem um feromônio para atrair o bicudo, e a armadilha tem formato cilíndrico. Os testes foram conduzidos pela empresa LiveFarm (ver www.livefarm.com.br) nos anos de 2021 e 2022. Foram obtidos 1764 registros de dados, considerando todas as armadilhas das fazendas citadas. A Figura 3 ilustra uma das armadilhas usadas.



Figura 3. Armadilha para bicudo da LiveFarm. Fonte: [LIVEFARM 2023]

Foram colhidos, em cada semana dos anos de 2021 e 2022, dados sobre a quantidade de bicudos dentro e fora da armadilha, e também os dados climáticos (temperatura e umidade do ar, radiação solar, precipitação, temperatura do ponto de orvalho, velocidade do vento) e latitude e longitude de cada armadilha. Inicialmente, o critério de classificação foi definido como risco alto e baixo e posteriormente, considerando n a quantidade de bicudos presentes em cada armadilha, foi definido da seguinte forma:

Se $n < 1 \Rightarrow$ o risco de aparecimento de bicudo é muito baixo

Se $n = 1 \Rightarrow$ o risco de aparecimento de bicudo é baixo

Se $1 < n \leq 2 \Rightarrow$ o risco de aparecimento de bicudo é médio

Se $2 < n \leq 4 \Rightarrow$ o risco de aparecimento de bicudo é alto

Se $x > 4 \Rightarrow$ o risco de aparecimento de bicudo é muito alto

Os dados obtidos em cada armadilha foram revistos quanto a erros, valores acima ou abaixo do intervalo aceitável, etc. Após uma limpeza nos dados, estes foram usados como entrada em algoritmos de machine learning (ML), para classificação, como random forest, xgboost, multi-layer perceptron (MLP), random tree, support vector machine, KNN, mapas auto-organizáveis (SOM), etc. Além dos algoritmos de machine learning, foram usados também modelos mistos, heurísticas e problema generalizado de atribuição. Esses algoritmos serão comparados quanto aos seus resultados obtidos e assim, o algoritmo que tiver melhor resultado (acurácia ou f1-score) será usado para gerar as predições. Não foram usadas series temporais pois existem alguns meses de 2021 e 2022 que não foram coletados dados e também pelo fato de ser apenas dois anos. Futuramente, com dados de 2023 em diante, serão testados algoritmos de series

temporais como Long Short-Term Memory (LSTM) [Hochreiter and Schmidhuber 1997], Gated Recurrent Unit [Cho et al., 2014], and Convolutional Neural Network [S. Albawi et al., 2017].

O software usado para executar os algoritmos de machine learning foi o WEKA (Waikato Environment for Knowledge Analysis), que pode ser visto em [WEKA 2023]. A Figura 4 ilustra a interface do Weka.

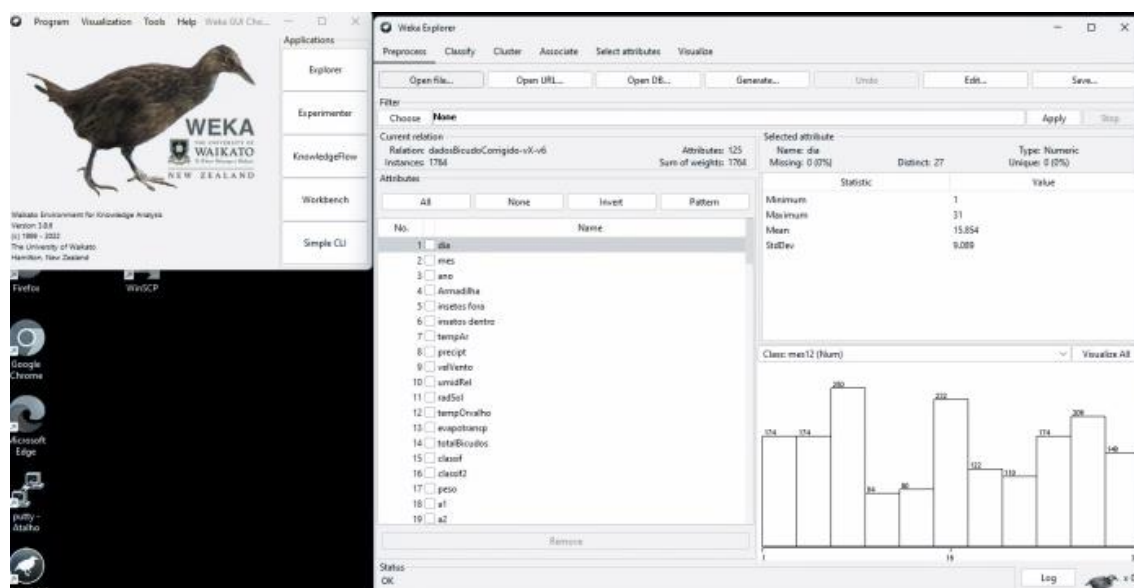


Figura 4. Página inicial do software Weka. Fonte: captura de tela feita pelos autores

O software weka permite a execução de uma série de algoritmos de machine learning classificadores ou regressores. Para cada algoritmo são fornecidas estatísticas como acurácia, f1-score, matriz de confusão, erro médio, etc. O uso deste software facilita a escolha do algoritmo a ser usado, visto que não precisa de programar os algoritmos. Além disso, o Weka permite parametrizar cada algoritmo e assim o usuário poderá inserir o número de épocas, a porcentagem de registros que devem ser usados para criar o modelo (treinamento), e depois testar e validar, se vai usar ou não validação cruzada, e parâmetros específicos para cada algoritmo a ser utilizado. Com os resultados obtidos de cada algoritmo, é possível escolher os melhores e então programar cada um dos algoritmos escolhidos usando Python ou R ou a linguagem mais conveniente e poder fazer alterações desejadas. Para o caso deste artigo, o melhor resultado de todas as análises, com diferentes algoritmos, a partir do qual foi gerado um modelo, foi codificado em python, para ser usado como biblioteca para outro software que necessitar de predição de bicudo, ou ainda, ser usado através do sitio <https://www.cnpaf.embrapa.br/predBicudo>, conforme pode ser visto na Figura 5. Este programa desenvolvido pelos autores será adicionado ao hardware da armadilha ilustrada na Figura 3, construída pela empresa Livefarm, citada anteriormente.

Acesso à Informação BRASIL

LIVEFARM Embrapa Arroz e Feijão

Sistema para prever a presença de bicudos na lavoura de algodão

Temperatura do ar: valores entre -20 e 50

Temperatura do ponto de orvalho: valores entre -20 e 50

Umidade relativa do ar: valores entre 0 e 100

Evapotranspiração: valores entre 0 e 100

Precipitação: valores entre 0 e 100

Radiação solar: valores entre 0 e 1000

Velocidade do vento: valores entre 0 e 20

Mês: --Escolha o mes--

ACESSAR

Figura 5 – Sistema de predição. Fonte: captura de tela feito pelos autores.

O software poderá ser usado também em ambiente Linux ou Windows, desde que tenha a linguagem python, versão 3.6, ou superior instalado. Inicialmente, o software foi pensado como integrante do software embarcado da armadilha automatizada para bicudos, ilustrado na Figura 3, mas com o tempo verificou-se que o software pode ser usado em outras situações. Este sistema descrito recebe como entrada um conjunto de variáveis climáticas (temperatura e umidade do ar, radiação solar, evapotranspiração, temperatura de orvalho, precipitação e velocidade do vento) e o mês que estas variáveis foram medidas e retorna como saída a predição do risco de aparecimento do bicudo na lavoura.

3. Resultados

As armadilhas foram dispostas ao redor da plantação (borda) e também no interior da plantação, conforme ilustrado na Figura 2. Dessa forma, os bicudos podem escolher qualquer armadilha para entrar. Os dados climáticos para as armadilhas, em cada uma das fazendas, eram praticamente os mesmos, dado que a área em que as armadilhas estavam não era tão grande, embora a distância entre elas variava entre 150 a 300 m. Como restrição deste problema, não deverá ser descartada nenhuma armadilha e seus resultados, mesmo que seja com a finalidade de melhoria do resultado final.

Para a classificação mais simples, como risco baixo ou alto, isto é, dois estados de risco, a acurácia e o fl score foram em torno de 85% com os algoritmos de machine learning. Porém, quando a classificação mudou para uma classificação mais rigorosa, como a citada acima (risco muito baixo, baixo, médio, alto e muito alto), os resultados mudaram, e o máximo obtido foi 80.4% com o algoritmo random Forest. O treinamento foi feito com 70% do conjunto de dados e a validação 30%. A Tabela 1 resume o resultado que foi obtido inicialmente.

Tabela 1. Algoritmos de ML testados e sua respectiva acurácia

Algoritmo	Acurácia (%)
Random Forest	80.4
J48	78.9
Classification Via Regression	78.8
REPTree	78.5
XGBoost Classifier	78.2
Multi Layer Perceptron (MLP)	78.1
Support Vector Machine (SVM)	76.9
JRIP	76.1
Random Tree	74.3
Decision Table	73.4

Foram testados também os algoritmos relativos ao problema generalizado de localização (PGA), que pode ser visto em [Lorena e Narciso 1999] e modelo linear misto, que pode se visto em [MLM 2023]. O PGA obteve 76.8% de acertos e o MLM obteve 74.8% de acertos.

O algoritmo random forest, ao ser executado com a opção de fornecer as variáveis que mais se destacaram, após 5 mil épocas, ou de maior importância, mostrou que as variáveis temperatura do ar, precipitação, velocidade do vento e umidade do ar são as mais importantes, nessa ordem, para esse conjunto de dados. As variáveis radiação solar, temperatura de orvalho e evapotranspiração são as que tem menor importância, mas ainda são significativas para o resultado final. No modelo gerado, optou-se por inserir o mês, visto que tem também forte correlação, conforme resultados obtidos quanto à importância de variáveis dado pelo algoritmo random forest. Desta forma, o melhor resultado de cada algoritmo da Tabela 1 acima ficou em função do mês, temperatura e umidade do ar, velocidade do vento, precipitação, temperatura de orvalho, evapotranspiração e radiação solar.

Observou-se que é possível contruir heurísticas para este problema, partindo do princípio de escolher cada uma das 5 possibilidades de classificação, de muito baixo a muito alto, para cada dia, de tal forma que maximize o índice de acerto. Uma heurística foi construída a partir do modelo gerado pelo random forest, e pode-se também obter um resultado levemente melhor do que o obtido pelos algoritmos testados. Entretanto, não é garantida a possibilidade de generalização do modelo, isto é, para outros dados que

forem colhidos daqui para frente, a heurística pode ter rendimento menor em termos de acertos. Desta forma, optou-se por usar o modelo gerado pelo random forest, que foi gerado para evitar erros de overfitting para assim contemplar generalização e ter bons resultados, mesmo com dados que não foram os de treinamento e validação.

Durante a execução dos algoritmos, observou-se que as armadilhas influenciam os resultados no sentido de melhorar. Isso foi constatado ao inserir as armadilhas como entrada de dados, através de código para cada armadilha, e o fato de ter ou não entrado bicudo em cada armadilha. Como isso é difícil generalizar, foi mantida a entrada de dados para o software desenvolvido, isto é, o mês e os dados climáticos. Por outro lado, existem fatores externos que influenciam a quantidade de bicudo que aparece na lavoura, como soqueiras e tigueras, matas próximas que abrigam bicudos, o que também não é fácil medir ou prever.

4. Conclusão

Este trabalho teve como objetivo mostrar uma maneira de ajudar o produtor a saber sobre possível aparecimento de bicudo na lavoura de algodão através de um software, a qualquer momento, e assim poder tomar decisão sobre o que fazer.

O tempo para o pragueiro ir até o local e fazer inspeções para dizer se tem ou não bicudo na lavoura e que decisão tomar pode levar alguns dias, como citado por [Rodrigues e Miranda 2015]. Este sistema, pode ser usado a qualquer tempo, o que é uma vantagem para ser usado até mesmo pelo pragueiro como ferramenta de auxílio a inspeção da lavoura em tempos oportunos.

O resultado obtido pelos algoritmos de ML podem ser melhorados à medida que mais dados forem sendo obtidos de novos experimentos que serão feitos a cada ano. Com um maior volume de dados, espera-se que os resultados sejam melhores do que os obtidos até aqui. Vale a pena mencionar que existe o tempo certo para a coleta de dados nos experimentos e estes têm um custo associado.

Referências Bibliográficas

AGRONOVAS, Controle de pragas no algodão. Sítio disponível em <http://www.agronovas.com.br/76209>. Acesso em 10/09/2023.

CANAL_RURAL. Bicudo do algodoeiro. Sítio disponível em <https://www.canalrural.com.br/projetos/sites-e-especiais/bicudo-do-algodoeiro-saiba-tudo-sobre-inseto-melhores-formas-controle-praga-9446/>, Acesso em 10/09/2023.

CHO, K., MERRIENBOER, B. V., GULCEHRE, C., BAHDANAU, D., BOUGARES, F., SCHWENK, H. AND BENGIO, Y. "Learning phrase representations using RNN encoder-decoder for statistical machine translation", Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, 2014.

- DEGRANDE, P.E. ; IDE, M.A.; NAKANO, O.; PARRA, J.R.P. Efeito de diversas temperaturas sobre a emergência do bicudo do algodoeiro (*Anthonomus grandis* Boheman, 1843) (Coleoptera-Curculionidae). In: CONGRESSO BRASILEIRO DE INICIAÇÃO CIENTÍFICA EM CIÊNCIAS AGRÁRIAS, 3., 1983, Florianópolis, SC. Resumos. Florianópolis, p.33, 1983
- GABRIEL, D. (2016). “O bicudo do algodoeiro”. Documento Técnico 25, Agência Paulista de Tecnologia dos Agronegócios, São Paulo.
- HOCHREITER, S. AND SCHMIDHUBER, J. “Long short-term memory”, Neural computation, MIT Press, 1997, vol. 9, number 8, pp. 1735–1780.
- LIVEFARM. LiveFARM Tecnologia Agropecuária Ltda. Sítio disponível em <<https://www.livefarm.com.br/>>. Acesso em 28/09/2023.
- MLM. Modelos Lineares Mistos. Sítio disponível em <http://labtrop.ib.usp.br/doku.php?id=cursos:planeco:roteiro:11-lmm> Acesso em 10/09/2023.
- NARCISO, M.G. AND LORENA, L.A.N. (1999) Lagrangean/surrogate relaxation for generalized assignment problems, European Journal of Operational Research, 114 165-177.
- NEGÓCIO_RURAL. Revista eletrônica “Negócio Rural”, Edição 24, páginas 12 e 13. Sítio disponível em <https://www.revistanegociorural.com.br>. Acesso em 10/09/2023.
- RODRIGUES, S.M.M.; MIRANDA, J.E. Controle etológico do bicudo-do-algodoeiro. In.: Instituto Mato-grossense do Algodão - IMAMT. O bicudo-do-algodoeiro (*Anthonomus grandis* Boh., 1843) nos cerrados brasileiros: Biologia e medidas de controle. Cuiabá, 2015. p.95-116.
- S. ALBAWI, S., MOHAMMED, T. A. AND AL-ZAWI, S. “Understanding of a convolutional neural network”, International Conference on Engineering and Technology, 2017, pp. 1–6.
- VOLPE, C.A.; LARA, F.M.; SOARES, J.J. Influência da temperatura e umidade do solo na mortalidade em *Anthonomus grandis* BOHEMAN, 1843 (Coleoptera, Curculionidae). In: CONGRESSO BRASILEIRO DE ENTOMOLOGIA, 14., 1993, Piracicaba, SP. Resumos. Piracicaba: SEB, 1993. p.38.
- WEKA. Machine learning. Sítio disponível em <<https://www.cs.waikato.ac.nz/ml/weka/>>. Acesso em 10/09/2023.