

Comparando o Desempenho de Algoritmos de Aprendizado de Máquina na Previsão da Cercosporiose do Café.

Eduardo Custodio Zanuni¹, Pedro Antonio Lorençone¹, João Antônio Lorençone¹, Michael Steinhorst Alcantara¹, Alisson Gaspar Chiquitto¹, Lucas Eduardo de Oliveira Aparecido², Guilherme Botega Torsoni¹

¹Instituto Federal de Mato Grosso do Sul– Naviraí-MS, ²IFSULDEMINAS– Muzambinho-MG

eduardo.zanuni@estudante.ifms.edu.br, guilherme.torsoni@ifms.edu.br

Resumo

O presente estudo visa prever a Cercosporiose em cafeeiros, uma doença fúngica causada pelo *Cercospora coffeicola*, utilizando algoritmos de machine learning (ML) baseados em dados agrometeorológicos. Dados meteorológicos e da doença foram obtidos em oito localidades de São Paulo e Minas Gerais entre 2010 e 2012, totalizando 191 epidemias. A análise considerou o período de latência da doença (7 dias) e variáveis climáticas como avgT, avgUR, entre outras. O XGBoost mostrou-se mais eficaz para cafeeiros com carga alta, apresentando $R^2=0.77$ e $RMSE=10,53$, enquanto o SVM foi superior para carga baixa com $R^2=0.76$ e $RMSE=12.82$. Observou-se maior incidência da doença em cafeeiros de carga alta. Regiões como Patrocínio-MG mostraram baixa incidência da doença. Conclui-se que, utilizando ML e variáveis agrometeorológicas, é possível estimar a incidência da Cercosporiose em cafeeiros.

Palavras-chave: Agroclimatologia, Fungos em café, Machine learning agrícola.

Introdução

A doença Cercosporiose, desencadeada pelo fungo *Cercospora coffeicola*, afeta folhas e frutos do cafeeiro durante todas as etapas de seu crescimento (VALE et al., 2021). Os sintomas associados a esta infecção incluem folhas tornando-se amareladas, queda prematura de folhas, e amadurecimento antecipado dos frutos (ANDRADE et al., 2021). Além disso, as lesões provocadas por esta doença tornam a planta vulnerável a outras infecções, comprometendo tanto a quantidade quanto a qualidade da produção de café (ESGARIO; KROHLING; VENTURA, 2020). Aspectos climáticos, como níveis de luz, temperatura, umidade atmosférica e duração da umidade sobre a planta, desempenham um papel crucial na manifestação desta doença (BOTELHO et al., 2017).

Nos tempos modernos, a tecnologia tem permitido o uso de ferramentas computacionais e estruturas matemáticas para antecipar a ocorrência de doenças nas plantações (CHEN et al., 2022). Técnicas de inteligência artificial, como redes neurais e florestas aleatórias, têm se revelado promissoras em previsões agronômicas (LEE et al., 2021). Muitas pesquisas têm se concentrado em usar esses métodos para modelar condições agrometeorológicas e antecipar a presença de pragas e enfermidades (GUO et al., 2021). Contudo, percebe-se uma escassez de estudos que explorem tais algoritmos para

a previsão da Cercosporiose em cafeeiros, particularmente no contexto brasileiro, um dos maiores produtores globais de café.

Dessa forma, objetivou-se utilizar variáveis agrometeorológicas juntamente com modelos de machine learning para espacializar e prever a incidência da cercosporiose do cafeeiro nas principais regiões produtoras de café no Brasil.

Metodologia

O trabalho foi executado em duas etapas. Primeiramente, foram utilizados dados meteorológicos e da doença do café obtidos em 8 localidades do estado de São Paulo e Minas Gerais, compreendendo regiões do sul e do cerrado mineiro. Nesta etapa foram calibrados diferentes tipos de ML e topologias para previsão da cercosporiose, com uma antecipadamente possível para uma tomada de decisão. As topologias com melhores performances foram utilizadas na segunda etapa do trabalho, que foi a previsão e espacialização da severidade de cercosporiose para todos os municípios relevantes para a produção de café no Brasil.

1. Coleta e Calibração dos Dados

Os algoritmos utilizados na previsão foram os seguintes: Support Vector Machine (SVM) e Extreme gradient boosting Regression (XGBoost).

Tabela 1. Parâmetros calibrados, os valores testados em cada parâmetro e os valores definidos em cada algoritmos na previsão da cercosporiose do cafeeiro para todo Brasil.

Modelo	Parameter	Defined values
SVM	kernel	rbf
	c	200
	epsilon	default
	gamma	auto
XGBoost	nthread	5
	learning_rate	0.3
	max_depth	5
	min_child_weight	5
	silent	1
	subsample	0.8
	colsample_bytree	0.03
n_estimators	1000	

2. Espacialização dos Dados

Com todos os algoritmos calibrados para a previsão da doença foi especializado as previsões do modelo mais acurado para as regiões com maior produção de café do Brasil, por meio do programa QGIS 3.32. O método de interpolação utilizado foi a krigagem (KRIGE, 1951), com o modelo esférico, um vizinho e resolução de 0,25° (25 km).

Resultados e Discussão

Espacialização dos Dados Carga Alta

Utilizando o modelo XGBoost e SVG foi possível estimar a incidência de brown eye em todas as localidades da região com maior produção de café arábica do país, cerca de 99.95 % do total. Foi perceptível diferença entre os cafeeiros em carga alta e carga baixa, sendo durante a primeira maior a severidade de brown eye na cultura (Figure 1 e 2). Quando o cafeeiro se encontra em carga alta ele direciona grande parte de seus foto assimilados para produção de frutos e enchimento de grãos, dificultando a manutenção de mecanismo de defesa, tornando-o mais suscetível (CAMARGO; CAMARGO, 2001).

Em carga alta para a região as classes mais predominantes foram de 0-10, 20-30 e 10-20%, médias anuais de 15.55, 36.11 e 52.63%, respectivamente. A classe mais grave de infestação (40-50%) concentrou em apenas 8.59% nos meses de maio, junho e julho. Durante junho 55.70% da área apresentou as duas classes mais graves, 33.48% para 30-40 e 13.58% em 40-50% (Figura F G). Por outro lado, outubro e novembro demonstraram maior concentração de áreas com severidade abaixo de 20%, média de 24.51 e 55.39% para 0-10 e 10-20% (Figure IJ K)

A ocorrência de brown eye variou conforme o estado dentro da região de estudo, com menores ocorrências nas localidades com maiores produções de café. Os únicos estados que apresentaram incidência dentre 40-50% foram BA (12.36%) e GO (12.52%), apresentaram a maior área com 20-30 também, 19.99 e 25.08%. Tanto na BA como em GO, junho e julho são os períodos mais críticos, em junho a média foi de 19.10% da área dos estados com cerca de 40 a 50% da lavoura contaminada. Ambos os estados demonstraram suscetibilidade ao ataque de *Cercospora coffeicola*, sendo necessário adotar medidas preventivas ao ataque do patógeno, métodos culturais como consórcio com árvores frutíferas, no qual o cafeeiro cresce sombreado, impossibilitando a ativação da toxina do fungo, não prejudicando as folhas (MOUEN BEDIMO et al., 2007).

Os municípios com maiores produções do país se encontram em regiões com baixa incidência da doença. Patrocínio – MG localizado a 426.7 km a noroeste de Belo Horizonte (Figure 1) é o maior produtor de café (1 milhão e

84 mil sacas year⁻¹), foi classificado durante todo o ano de carga alta abaixo de 20% de incidência, em geral as lavouras apresentaram menos de 10%. Da mesma maneira, para os estados de SP, ES, BA, PR e GO os municípios com maiores produções: Pedregulho, Iúna, Barra do Choça, Carlópolis, Porciúncula e Cristalina, respectivamente, apresentaram incidência semelhantes a Patrocínio. Essas localidades alcançam maiores produtividades devido não sofrerem com altas perdas de área foliar devido a brown eye, nem com perdas de grãos que serão colhidos.

Algumas regiões apresentaram condições de maior risco do aparecimento de doença, principalmente ao norte de MG, de GO e da BA para cafeeiros em carga alta (Figure 1). Localidades como Monte Carlos ao norte de MG, Campo Formoso norte da BA, Alto Paraíso de Goiás demonstraram incidência acima de 30% da lavoura de café, durante maio, junho e julho, chegando até 50% das plantas infectadas. Cafeeiros cultivados nessa região chegam na maturação (maio e junho) com grande parte dos grãos perdidos devido necrose causada pelo fungo, além de prejudicar a maturação dos frutos restantes devido a alta taxa de desfolha (COSTA et al., 2014). Dessa forma, a maior parte das localidades que apresentaram incidência elevada de *Cercospora coffeicola*, é necessário medidas preventivas como a aplicação foliar de produtos à base de cobre (PATRÍCIO et al., 2008).

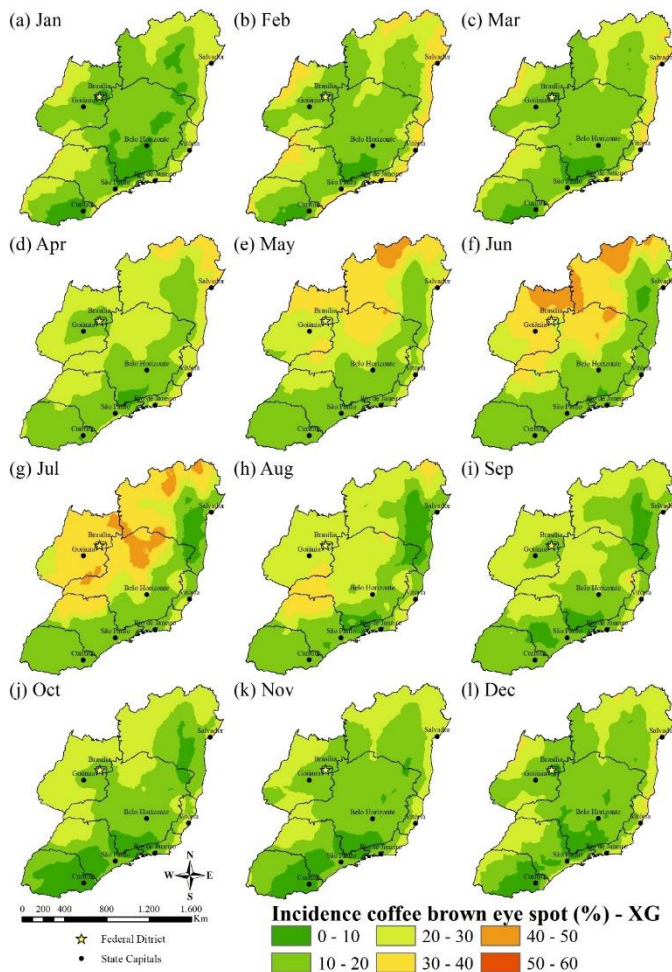


Figura 1. Incidência mensal de cercosporiose do café em carga alta previsto pelo algoritmo XG para todas as regiões produtoras do Brasil.

Espacialização dos Dados Carga Alta

O modelo mais acurado para prever a cercosporiose em café em carga baixa foi o SVM. Dentro das classes de incidência de cercosporiose as que apresentaram valores na região estudada foram, 0-10 com 65.52 %, 10-20 com 56.33 % e 20-30 com 4.07 % (Figura 2). Em relação a classe de 0-10 o estado que apresentou maior área foi o Espírito Santo com 84.13 % de seu território durante todo ano, já o estado com a menor porcentagem foi Bahia com 43.67 %. Para a classe 10-20 o estado do Rio de Janeiro é o que apresentou a maior porcentagem 64.11 % de seu território, sendo o Minas Gerais o estado com menor porcentagem 41.45 %. Já para a classe 20-30 apenas alguns estados apresentaram porcentagem, sendo o Espírito Santo o maior, com 7.23 % (Figura 2).

Em relação a distribuição temporal do modelo SVM houve uma variação em relação aos meses do ano e as classes (Figura 2). Os meses de junho, julho e dezembro foram que apresentaram maior probabilidade de incidência de

cercosporiose na classe 0-10 com valores de 95.85, 94.16 e 86.69 % respectivamente. Para a classe 10-20 os meses de abril, outubro e setembro apresentaram as maiores incidências de cercosporiose com 98.11, 91.36 e 89.49 % respectivamente, sendo o mês de abril um período importante para formação do cafeeiro em que ocorre a maturação das gemas, principalmente em anos de carga baixa, afetando dessa forma diretamente a produção da cultura (CAMARGO; CAMARGO, 2001).

Na região Sul de Minas Gerais, a classe mais predominante de incidência de cercosporiose para cafeeiros em carga baixa foi 0-10 com 61,78%. A distribuição temporal nessa região se concentrou principalmente nos meses de dezembro, janeiro e fevereiro.

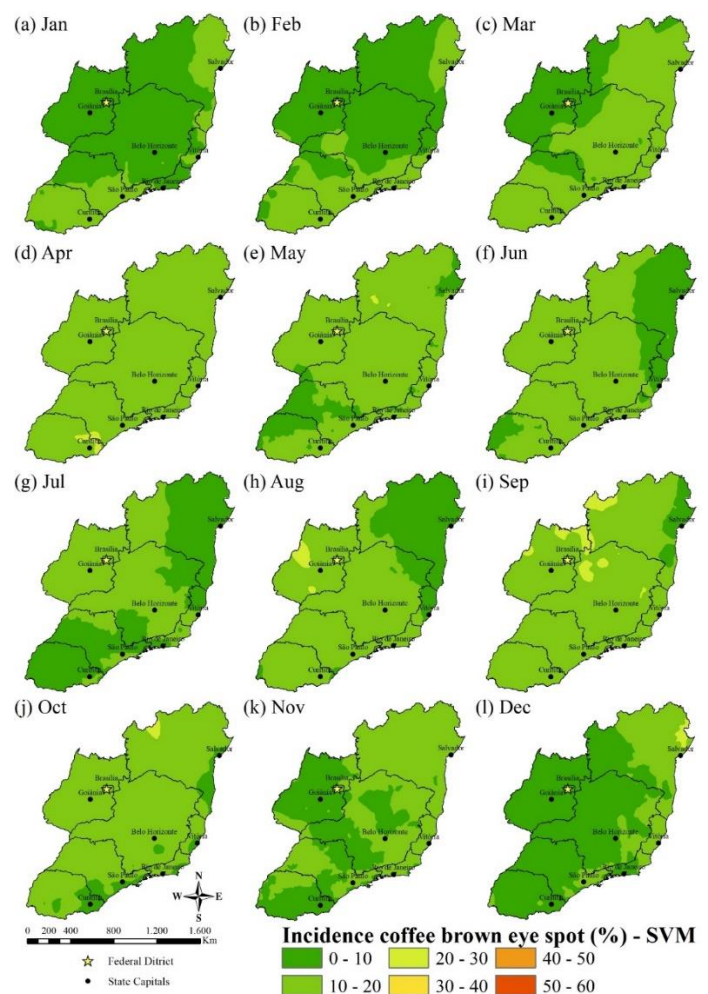


Figura 2. Incidência mensal de cercosporiose do café em carga baixa previsto pelo algoritmo SVM para todas as regiões produtoras do Brasil.

Considerações Finais

Utilizando o modelo XGBoost e SVG é possível estimar a incidência de brown eye em todas as localidades da região com maior produção de café arábica do país com 7 dias de antecedência.

Os cafeeiros em carga alta apresentam uma incidência maior em relação em carga baixa, com valores médios para alta de 36.11% e para baixa de 4.07% para classe de 20-30%, já para 0-10% as médias foram 15.55% e 65.52%, respectivamente.

As localidades com maiores produções de café apresentam baixa incidência de brown eye spot durante o ano todo, tanto para carga alta e baixa. Em carga alta o norte de Minas Gerais, da Bahia e Nordeste de Goiás demonstram alta ocorrência da doença durante maio a julho, alcançando 50% da lavoura contaminada em junho.

Agradecimentos

Agradecemos ao IFMS e o CNPq pelo auxílio financeiro e apoio para execução desse trabalho..

Referências

ALLEN, R. G. et al. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. **Fao, Rome**, v. 300, n. 9, p. D05109, 1998.

ANDRADE, C. C. L. et al. Infection process and defense response of two distinct symptoms of *Cercospora* leaf spot in coffee leaves. **Phytoparasitica**, v. 49, n. 4, p. 727–737, set. 2021.

APARECIDO, L. E. DE O. [UNESP; ROLIM, G. DE S. [UNESP; DE SOUZA, P. S. Sensitivity of newly transplanted coffee plants to climatic conditions at altitudes of Minas Gerais, Brazil. **Australian Journal of Crop Science**, p. 160–167, 1 jan. 2015.

BOTELHO, D. M. DOS S. et al. Cercosporiosis resistance in coffee germplasm collection. **Euphytica**, v. 213, n. 6, p. 117, jun. 2017.

CAMARGO, Â. P. D.; CAMARGO, M. B. P. D. Definição e esquematização das fases fenológicas do cafeeiro arábica nas condições tropicais do Brasil. **Bragantia**, v. 60, n. 1, p. 65–68, 2001.

CHALFOUN, N. **Design and application of natural down-draft evaporative cooling devices**. . Em: PROCEEDINGS OF THE NATIONAL PASSIVE SOLAR CONFERENCE. AMERICAN SOLAR ENERGY SOCIETY INC, 1997.

CHEN, Z. et al. Plant Disease Recognition Model Based on Improved YOLOv5. **Agronomy**, v. 12, n. 2, p. 365, 31 jan. 2022.

CORNELL, J.; BERGER, R. Factors that influence the value of the coefficient of determination in simple linear and nonlinear regression models. **Phytopathology**, v. 77, n. 1, p. 63–70, 1987.

COSTA, B. H. G. et al. Suppression of Rust and Brown Eye Spot Diseases on Coffee by Phosphites and By-products of Coffee and Citrus Industries. **Journal of Phytopathology**, v. 162, n. 10, p. 635–642, out. 2014.

ESGARIO, J. G. M.; KROHLING, R. A.; VENTURA, J. A. Deep learning for classification and severity estimation of coffee leaf biotic stress. **Computers and Electronics in Agriculture**, v. 169, p. 105162, fev. 2020.

GUO, Y. et al. Integrated phenology and climate in rice yields prediction using machine learning methods. **Ecological Indicators**, v. 120, p. 106935, jan. 2021.

KRIGE, D. G. A statistical approach to some basic mine valuation problems on the Witwatersrand. **Journal of the Southern African Institute of Mining and Metallurgy**, v. 52, n. 6, p. 119–139, 1951.

KUSHALAPPA, A.; AKUTSU, M.; LUDWIG, A. Application of survival ratio for monocyclic process of *Hemileia vastatrix* in predicting coffee rust infection rates. **Phytopathology**, v. 73, n. 1, p. 96–103, 1983.

LEE, J. et al. Evaluation of Rainfall Erosivity Factor Estimation Using Machine and Deep Learning Models. **Water**, v. 13, n. 3, p. 382, 1 fev. 2021.

MONTEITH, J. L. Dew. **Quarterly Journal of the Royal Meteorological Society**, v. 83, n. 357, p. 322–341, jul. 1957.

PATRÍCIO, F. R. A. et al. Effectiveness of acibenzolar-S-methyl, fungicides and antibiotics for the control of brown eye spot, bacterial blight, brown leaf spot and coffee rust in coffee. **Annals of Applied Biology**, v. 152, n. 1, p. 29–39, fev. 2008.

SPARKS, A. H. nasapower: a NASA POWER global meteorology, surface solar energy and climatology data client for R. **Journal of Open Source Software**, v. 3, n. 30, p. 1035, 2018.

STACKHOUSE, P. W. et al. Prediction of Worldwide Energy Resource (POWER)-Agroclimatology methodology-(1.0 latitude by 1.0 longitude spatial resolution). **Hampton, NASA Langley Research Center**, 2015.

THORNTON, C.; MATHER, J. The water balance publications in Climatology, 8 (1). **DIT, Laboratory of climatology, Centerton, NJ, USA**, 1955.

VALE, P. A. S. et al. Epitypification of *Cercospora coffeicola* and its involvement with two different symptoms on coffee leaves in Brazil. **European Journal of Plant Pathology**, v. 159, n. 2, p. 399–408, fev. 2021.