

UNIVERSIDADE ESTADUAL PAULISTA
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CÂMPUS DE JABOTICABAL

MODELOS AGROMETEOROLÓGICOS-ESPECTRAIS BASEADOS EM
INTELIGÊNCIA ARTIFICIAL PARA ESTIMATIVA DA SEVERIDADE DA
CERCOSPORIOSE NO CAFEEIRO.

Bruna Bueno Abrahão

Jaboticabal – SP

UNIVERSIDADE ESTADUAL PAULISTA
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CÂMPUS DE JABOTICABAL

MODELOS AGROMETEOROLÓGICOS-ESPECTRAIS BASEADOS EM
INTELIGÊNCIA ARTIFICIAL PARA ESTIMATIVA DA SEVERIDADE DA
CERCOSPORIOSE NO CAFEEIRO.

Bruna Bueno Abrahão

Orientador: Prof. Dr. Glauco de Souza Rolim

Trabalho apresentado à Faculdade de Ciências
Agrárias e Veterinárias – UNESP, Câmpus de
Jaboticabal, para a graduação em
ENGENHARIA AGRONÔMICA

Jaboticabal – SP
2º Semestre/2024

A159m

Abrahão, Bruna Bueno

Modelos agrometeorológicos-espectrais baseados em inteligência artificial para estimativa da severidade da cercosporiose no cafeeiro. / Bruna Bueno Abrahão. -- Jaboticabal, 2024

39 p. : il., tabs., fotos

Trabalho de conclusão de curso (Bacharelado - Engenharia Agrônômica) - Universidade Estadual Paulista (UNESP), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal

Orientador: Glauco de Souza Rolim

1. Variáveis climáticas. 2. Cercosporiose. 3. Random Forest. 4. Regressão Linear Múltipla. 5. Método da árvore de decisão. I. Título.



UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
Câmpus Jaboticabal



BRUNA BUENO ABRAHÃO

**MODELOS AGROMETEOROLÓGICOS-ESPECTRAIS BASEADOS EM INTELIGÊNCIA
ARTIFICIAL PARA ESTIMATIVA DA SEVERIDADE DA CERCOSPORIOSE NO
CAFEIEIRO**

Trabalho de Conclusão de Curso apresentada à Universidade Estadual Paulista (UNESP), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal, para obtenção do título de Bacharel em Engenharia Agrônoma.

Orientador: Prof. Dr. Glauco de Souza Rolim

Área de Concentração: Agrometeorologia

Data da defesa: 06 / 11 / 2024

(x) Aprovado

() Reprovado

Banca Examinadora:

Prof. Dr. Glauco de Souza Rolim

UNESP – Faculdade de Ciências Agrárias e Veterinárias – Câmpus de Jaboticabal

Andrea Onelia Rodriguez Roa

UNESP – Faculdade de Ciências Agrárias e Veterinárias – Câmpus de Jaboticabal

Thaís Rayane Gomes da Silva

UNESP – Faculdade de Ciências Agrárias e Veterinárias – Câmpus de Jaboticabal

Aprovado por ato “ad referendum” do presidente do Conselho do Departamento em: 13-11-2024.

Prof. Dr. Danísio Prado Munari
Chefe do Departamento

AGRADECIMENTOS

Agradeço aos meus familiares, especialmente meus pais, que sempre me incentivaram e me apoiaram.

Ao meu namorado, Marco Aurélio Gabas, por todos os momentos que vivemos juntos ao longo dos últimos cinco anos e todo o apoio e carinho.

A todas as moradoras da República Dama de Copos, pela amizade de todos esses anos.

Às minhas amigas Ana Júlia, Maria Gabriela e Isabela, por terem sido meu porto seguro, durante os cinco anos do curso e que levo no meu coração para a vida toda.

Ao meu professor orientador, Glauco de Souza Rolim, pela orientação, pela paciência e pelos ensinamentos durante toda esta jornada.

A Adriana Takakura, Secretária do Departamento de Ciências Exatas, por todo apoio.

À Universidade Estadual Paulista, especialmente, ao grupo GAS- Group of Agrometeorological Studies.

Ao Conselho Nacional de Desenvolvimento Científico – CNPQ, pelo auxílio financeiro que me ajudou no desenvolvimento desta pesquisa.

E por fim, aos membros da banca de defesa, pela disponibilidade e pelos ensinamentos valiosos.

Muito obrigada.

SUMÁRIO

RESUMO	vii
ABSTRACT	viii
1. INTRODUÇÃO	9
2. MATERIAL E MÉTODOS	15
2.1. Dados Meteorológicos	15
2.2. Índice de Vegetação	18
2.3. Dados da Cercosporiose	20
2.4. Modelos de Estimativas	21
3. RESULTADOS	24
4. CONCLUSÃO	31
5. REFERÊNCIAS	32

RESUMO

A produção de café é considerada uma atividade extremamente importante para a economia brasileira, sendo uma das bebidas mais consumidas pela população mundial, perdendo somente para a água. O padrão de qualidade do café é diretamente ligado ao clima, influenciando na produtividade da cultura e na incidência e severidade de doenças, como a cercosporiose. Este trabalho tem o objetivo de realizar a estimativa da severidade da cercosporiose para as cidades de Mococa, Franca, Marília, Garça, Ourinhos, Avaré, São João da Boa Vista e Varginha, com base em modelos agrometeorológicos-espectrais, do período de 2015 a 2022. Os dados meteorológicos estudados incluem precipitação, umidade relativa, temperatura (média, máxima e mínima), velocidade do vento, saldo de radiação umidade do solo e temperatura no ponto de orvalho, os quais foram obtidos a partir da plataforma NASA-POWER. Os índices de NDVI e EVI foram obtidos através da plataforma SATveg, fornecida pela Embrapa. Os dados sobre a severidade da doença cercosporiose foram extraídos do sistema da Fundação Procafé. Os modelos de estimativas utilizados são Regressão Linear Múltipla (RLM), Árvore de decisão (AD) e Random Forest (RF). Os Resultados confirmam o impacto do clima sobre a severidade da cercosporiose no cafeeiro, e mostram que modelos de estimativa como Random Forest e Árvore de Decisão tem um desempenho maior se comparado à Regressão Linear Múltipla.

Palavras-chave: cercosporiose, variáveis climáticas, Random forest, regressão, Árvore de Decisão.

ABSTRACT

Coffee production is considered an extremely important activity for the Brazilian economy and is one of the most consumed beverages worldwide, second only to water. The quality of coffee is directly related to the climate, which influences both the productivity of the crop and the incidence and severity of diseases such as cercospora leaf spot. This study aims to estimate the severity of cercospora leaf spot for the cities of Mococa, Franca, Marília, Garça, Ourinhos, Avaré, São João da Boa Vista, and Varginha, using agrometeorological-spectral models from 2015 to 2022. The meteorological data studied include precipitation, relative humidity, temperature (mean, maximum, and minimum), wind speed, radiation balance, soil moisture, and dew point temperature, all of which were obtained from the NASA-POWER platform. NDVI and EVI indices were acquired through the SATveg platform provided by Embrapa. Data on the severity of cercospora leaf spot were extracted from the Fundação Procafé system. The estimation models used are Multiple Linear Regression (MLR), Decision Tree (DT), and Random Forest (RF). The results confirm the impact of climate on the severity of cercospora leaf spot in coffee and show that estimation models such as Random Forest and Decision Tree perform better compared to Multiple Linear Regression.

Keywords: cercosporiosis, climatic variables, Random Forest, regressor, Decision Tree.

1. INTRODUÇÃO

A produção de café desempenha um papel fundamental na economia brasileira, sendo uns dos principais produtos de exportação do país e a segunda bebida mais consumida do mundo, perdendo somente para a água (ZELBER-SAGI et al., 2015). O Brasil é o maior produtor e exportador mundial de café, produzindo na safra de 2020, 63 milhões de sacas beneficiadas, enquanto na safra de 2021 apresentou uma queda de 21%, totalizando 48,80 milhões de sacas beneficiadas, resultado atribuído à bienalidade negativa da cultura (CONAB, 2020; 2021). Com mais de 2,18 milhões de hectares cultivados, Minas Gerais se destaca como o principal produtor do país, responsável por 55% da produção nacional, predominando o café arábica, enquanto o estado de São Paulo concentra cerca de 10 a 15% da produção (CONAB, 2021).

A qualidade da produção de café é altamente dependente de práticas culturais eficientes, assim como os fatores fisiológicos e ambientais, que desempenham um papel crucial no crescimento e desenvolvimento da cultura do café. Fatores como a temperatura e a disponibilidade hídrica, particularmente em estádios fenológicos críticos, afetam diretamente a produtividade do café (ASSAD et al., 2004; PICINI et al., 1999). As condições ideais de cultivo incluem chuvas bem distribuídas, ausência de déficit hídrico acima de 150 mm e temperaturas variando entre 18°C e 23°C, embora o cafeeiro possa tolerar uma ampla faixa de precipitação (ALÈGRE, 1959; DAMATTA & RENA, 2002; CARVALHO, 2008). Déficits hídricos severos, especialmente quando

combinados com temperaturas elevadas, podem reduzir a frutificação, resultando em menor produtividade (MATIELLO et al., 1991).

A produtividade do cafeeiro é diretamente influenciada pela ocorrência de doenças, que, por sua vez, são fortemente afetadas pelas condições climáticas (HARVEY et al., 2018; ROLIM et al., 2020). A distribuição da chuva ao longo do ano é um fator essencial na regulação da fenologia da planta e no aumento da severidade das doenças (CAMARGO, 2010). A temperatura do ar também é um elemento crítico, regulando tanto o desenvolvimento do cafeeiro quanto a intensidade das doenças (HODGES, 1990). Em particular, condições térmicas que variam de 5 °C a 30 °C, combinadas com alta umidade relativa e longos períodos de molhamento foliar, criam condições ambientes favoráveis para o desenvolvimento de doenças fúngicas (SENTELHAS et al., 2008; BERUSKI et al., 2015).

Outros fatores climáticos também afetam o desenvolvimento das doenças fúngicas, como a chuva, o vento, o fotoperíodo e a radiação solar, esses elementos influenciam a disseminação, a germinação e penetração de esporos (SENTELHAS et al., 2001).

A cercosporiose, ou mancha-de-olho pardo é uma das doenças mais relevantes que impacta a produção de café. Essa patologia é provocada pelo fungo *Cercospora coffeicola* (Berk e Cooke), pertence à família Dematiaceae, ordem Moniliales, classe dos fungos mitospóricos (GODOY et al., 1997). Afetam principalmente as folhas e os frutos do café, atacando desde mudas a até plantas adultas (LIMA et al., 2012). A cercosporiose pode ocasionar grandes perdas na produção de café, cerca de 30%, além disso causar danos à qualidade da bebida

(POZZA et al., 2001). A doença fúngica está disseminada em todas as regiões produtoras de café e demonstra alta agressividade nos cafezais (ZAMBOLIM et al., 2005).

O patógeno da cercosporiose ataca inicialmente as folhas do café, adentrando entre as aberturas na face superior da folha ou na cutícula dos frutos (SOUZA et al., 2012), atacando todas as fases de desenvolvimento do cafeeiro, do viveiro até o campo (ACEVES-NAVARRO et al., 2020). As folhas atacadas pelo fungo, caem, causando desfolha e seca dos ramos. Já nos frutos, as lesões vão aumentando, permanecendo até o amadurecimento. As manchas são caracterizadas por ter uma cor marrom-claro e arroxeadas, tendo maior incidência nos ramos expostos a luz solar (MATIELLO et al., 2010).

O clima interfere no desenvolvimento da doença, sendo esta favorecida pela luminosidade e pela radiação solar, de acordo com um estudo feito por Salgado et al. (2007), a radiação solar direta nas folhas do café, é um fator determinante para o desenvolvimento da cercosporiose, dessa forma, mudas que são mantidas a pleno sol são mais favoráveis ao surgimento da doença (ZAMBOLIM et al., 2005).

Outras variáveis climáticas que influenciam o desenvolvimento da doença são, temperatura média do ar de 10 a 25°C, umidade relativa maior que 90% (BOTELHO et al., 2017) e o período de água livre sobre o tecido vegetal, maior que 9h por dia (BOTELHO et al., 2017; SILVA et al., 2018). Além das condições ambientais, o nutricional da planta também influencia na intensidade da doença (CARVALHO et al., 2008), lavouras com deficiência hídrica prolongada, nutrição deficiente e desequilibrada apresentam maior progresso da doença no campo.

Nesse caso, a produção é reduzida e a infecção direta nos frutos afeta negativamente a qualidade do café.

O manejo da cercosporiose é feito de maneira preventiva desde os viveiros, na formação das mudas, utilizando substratos adubados adequadamente. Muitas vezes é utilizado o controle químico nos viveiros, principalmente antes da aclimação das mudas que serão levadas ao campo. Após as mudas serem transferidas para o campo, é recomendado evitar solos arenosos e controlar as adubações (ZAMBOLIM et al., 2005), porém a utilização de fungicidas para o controle da cercosporiose é, atualmente, a forma de manejo mais usual nos programas de tratamento fitossanitário na cultura do café (THOMAZIELLO et al., 2000).

Outra forma utilizada para fazer o controle preventivo da cercosporiose, é utilizando o sensoriamento remoto, de acordo com, Katsuhama et al. (2018) foram capazes de identificar áreas cultivadas com café atingidas por ferrugem com sucesso, utilizando imagens NDVI.

A identificação de doenças em cultivos, como no café, através do sensoriamento remoto é baseada em mudanças na resposta espectral das folhas causadas por danos estruturais (MARTINS et al., 2017). Patógenos podem alterar a reflectância, afetando bandas visíveis (VIS) e no infravermelho próximo (NIR), além de modificar a densidade do dossel e a área foliar da planta (MALTHUS et al., 1993; WEST et al., 2003). O índice de vegetação por diferença normalizada (NDVI) é usado para avaliar a desfolhação, captando doenças nas áreas agrícolas (MIRIK et al., 2006; KUCKENBERG et al., 2009). Estudos como

o de Katsuhama et al. (2018) mostram que índices de vegetação ajudam a monitorar doenças e otimizar o manejo na aplicação de fungicidas.

O café é a cultura que mais utiliza grandes quantidades de defensivos químicos no controle de pragas e doenças, atualmente, com alto perigo de contaminação ao meio ambiente. Os gastos com o controle de doenças fúngicas em uma lavoura de café adensado, do primeiro ao sexto ano, representam cerca de 36,86% dos gastos com os insumos e materiais de consumo (AGRIANUAL, 1999). Com isso, existem alternativas que devem ser exploradas para diminuir a dependência desse método, como a implementação da modelagem agrometeorológica, na qual permite simular a variabilidade de pragas e doenças da cafeicultura em relação aos fatores meteorológicos (ROLIM et al. 2008), evitando a pulverização excessiva de defensivos agrícolas, desse modo, trazendo benefícios ao meio ambiente, e consequentemente, diminuindo os custos com a produção (GILLESPIE et al., 2008).

Sistemas de aprendizagem de máquina, que realizam previsões de incidências de doenças, como na cultura do café, com base nas condições meteorológicas, auxiliam na tomada de decisões para a aplicação de defensivos agrícolas. Esses sistemas indicam os períodos mais propícios para o desenvolvimento do patógeno (XU, 2003).

Aprendizagem de máquinas é um campo da ciência que lida com identificação de padrões por meio da teoria da aprendizagem computacional dentro da inteligência artificial (SAHOO et al., 2019). Esses métodos são altamente eficazes para realizar análises de grande escala em um curto espaço de tempo, oferecendo grande eficácia e acurácia (LI et al., 2019).

No entanto, poucas pesquisas têm explorado a combinação de dados meteorológicos e espectrais para prever ou estimar doenças no cafeeiro, e a dificuldade de acesso às variáveis climáticas associadas ao surgimento de doenças, tem dificultado o manejo efetivo (POZZA et al., 2008). Dessa forma, a implementação de sistemas de alertas fitossanitários fundamentados em modelos de previsão de doenças, se apresenta como uma alternativa essencial tanto do ponto de vista econômico quanto ambiental.

Tendo em vista o que foi apresentado, o objetivo desse trabalho foi desenvolver um modelo agrometeorológico-espectral para estimar a severidade da cercosporiose na cultura do café para as regiões da Alta Mogiana, entre 2015 e 2022, utilizando Regressão Linear Múltipla, Árvore de decisão e aprendizagem de máquina pelo modelo Random Forest.

2. MATERIAL E MÉTODOS

2.1. Dados Meteorológicos

Os dados meteorológicos diários, incluindo temperatura média (T2M), máxima (Tmáx) e mínima do ar (Tmin), precipitação (P), umidade relativa (RH2M), saldo de radiação (Q0 e Qg), temperatura no ponto de orvalho (TPO), velocidade do vento (U2 ou WS2M) e umidade do solo (usolo), foram coletados através da plataforma NASA-POWER, do ano de 2015 a 2022 (<https://power.larc.nasa.gov>). Os locais analisados foram: Franca, Mococa, Marília, Garça, Ourinhos, Avaré, São João da Boa Vista e Varginha.

Os dados meteorológicos diários considerados foram referentes a um período 14 dias que antecede o final de cada mês (Tabela 1). Esse intervalo foi escolhido porque, ao término de cada mês, são divulgados os dados de severidade da Cercosporiose nos Boletins da Fundação Procafé, correspondendo ao período de incubação da doença nas regiões.

O sistema NASA-Power utiliza dados derivados do modelo de assimilação GMAO MERRA-2 e GEOS 5.12.4 FP-IT da NASA. O MERRA-2 é uma nova versão mais recente do sistema de assimilação de dados de observação da terra Goddard (GEOS) da NASA, enquanto o GEOS 5.12.4 tem a mesma resolução de grade do MERRA-2. As informações fornecidas pelo sistema NASA-Power Release-8 estão em uma grade global de resolução espacial de 0,5 ° x 0,5 °. A versão 8.0 do POWER tem o objetivo de disseminar dados meteorológicos voltados para o uso da captação de energia solar, construção sustentável utilizando energia fotovoltaica, e no campo da

Agroclimatologia, fornecendo dados compatíveis com o GIS e aplicativos da web habilitados para o GIS (NASA POWER, 2021).

A partir dos dados meteorológicos obtidos, a evapotranspiração de referência (ET_o) foi calculada utilizando o método de Penman-Monteith (FAO), conforme descrito na equação 1 (ALLEN et al., 1998).

:

$$ET_o = \frac{0,408\Delta(Rn - G) + \gamma \frac{900}{T + 273} U_2(e_s - e_a)}{\Delta + \gamma(1 + 0,34U_2)} \quad (Eq. 1)$$

em que, ET_o (PET) – evapotranspiração de referência (mm dia⁻¹), R_n – saldo de radiação à superfície do cultivo (MJ m⁻² dia⁻¹), G – densidade do fluxo de calor do solo (MJ m⁻² dia⁻¹), T – temperatura do ar a 2 m de altura (°C), U₂ – velocidade de vento a 2 m de altura (m s⁻¹), e_s – pressão de vapor de saturação (kPa), e_a – pressão parcial de vapor (kPa), Δ – declividade da curva de pressão de vapor de saturação (kPa), γ – coeficiente psicrométrico (kPa °C⁻¹).

A evapotranspiração real (ETR, mm), deficiência hídrica (DEF, mm) e excedente hídrico (EXC, mm) foram calculados com base no modelo sequencial de balanço hídrico proposto por Thornthwaite e Mather (1955), conforme ilustrado na Figura 1.

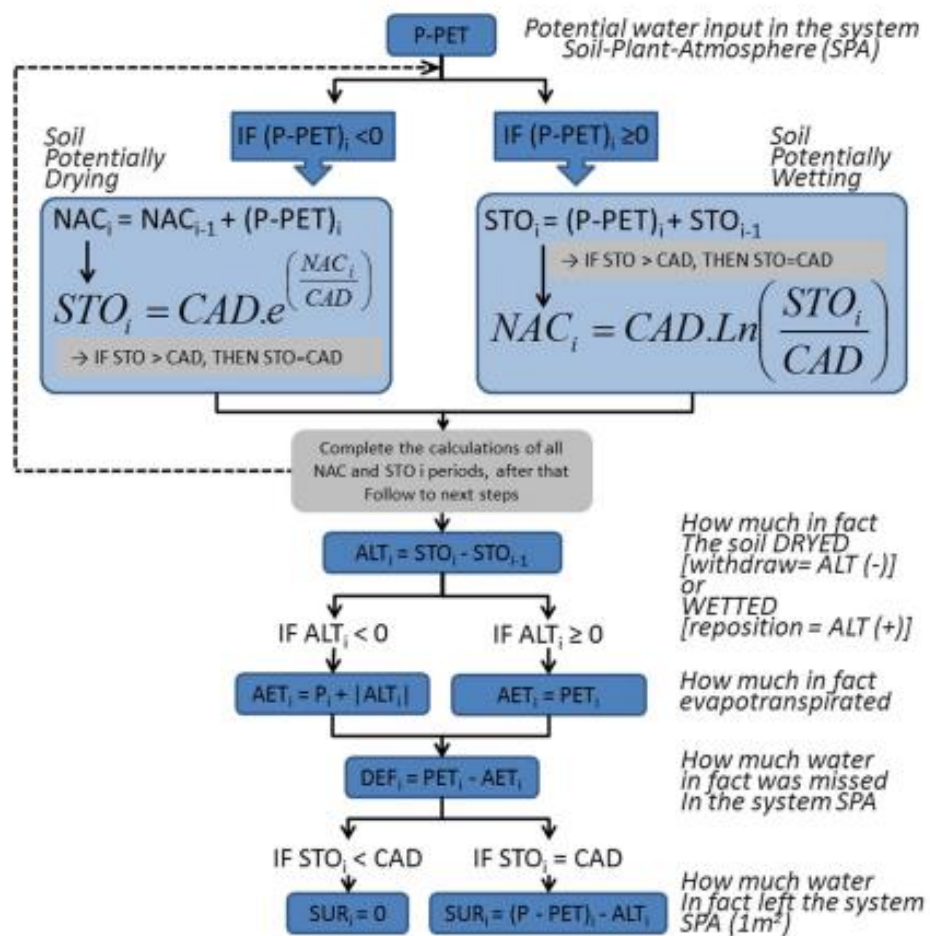


Figura 1. Fluxograma do modelo de balanço hídrico modificado de Thornthwaite e Mather (1955). Adaptado de Rolim et al. (2020).

Tabela 1. Variáveis utilizadas nos modelos ajustados

Variável	Dia 1	Dia 2	Dia 3	...	Dia 14
T2M (°C)	TM2_1	TM2_2	TM2_3	...	TM2_14
TPO (°C)	TPO_1	TPO_2	TPO_3	...	TPO_14
Tmin (°C)	Tmin_1	Tmin_2	Tmin_3	...	Tmin_14
Tmáx (°C)	Tmax_1	Tmax_2	Tmax_3	...	Tmax_14
RH2M (%)	RH2M_1	RH2M_2	RH2M_3	...	RH2M_14
P (mm/dia)	P_1	P_2	P_3	...	P_14
Qg (MJ/m2/dia)	Qg_1	Qg_2	Qg_3	...	Qg_14

Q0 (MJ/m2/dia)	Q0_1	Q0_2	Q0_3	...	Q0_14
ETo (mm/dia)	PET_1	PET_2	PET_3	...	PET_14
ETR (mm)	ETR_1	ETR_2	ETR_3	...	ETR_14
Usolo (%)	usolo_1	usolo_2	usolo_3	...	usolo_14
DEF (mm)	DEF_1	DEF_2	DEF_3	...	DEF_14
EXC (mm)	EXC_1	EXC_2	EXC_3	...	EXC_14
U2 (m/s)	U2_1	U2_2	U2_3	...	U2_14

2.2. Índice de Vegetação

Os índices de vegetação com diferença normalizada (NDVI, conforme a equação 2) e o índice de vegetação melhorado (EVI, conforme a equação 3) foram coletados através da plataforma SATveg (Sistema de Análise Temporal da Vegetação).

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (Eq. 2)$$

em que, NIR é a reflectância no comprimento de onda infravermelho próximo e Red é a reflectância média do vermelho.

$$EVI = G \times \frac{NIR - Red}{NIR + C1 \times RED - C2 \times Blue + L} \quad (Eq. 3)$$

em que, Blue é a reflectância no comprimento de onda médio no azul; C1, C2, L e G são coeficientes de ajuste relacionados à quantidade de aerossóis iguais a 6, 7.5, 1 e 2.5, respectivamente.

O SATveg é uma plataforma que pode ser acessada de forma online, foi desenvolvida pela Embrapa Informática Agropecuária, e permite acessar e

visualizar perfis temporais do índice vegetativo de NDVI e EVI a partir do sensor MODIS em diferentes locais da América do Sul.

Para cada região, foram selecionados aproximadamente 10 a 20 polígonos em áreas reconhecidamente com café, foi encontrado um padrão para reconhecer e diferenciar área com cultivo de café e cultivo de citros. As médias de NDVI e EVI foram calculadas para representar cada uma das regiões, conforme ilustrado na Figura 2.

Os dados dos índices de vegetação são disponibilizados a cada 16 dias a partir do satélite TERRA. Assim os valores de NDVI e EVI foram sincronizados ao mesmo mês dos dados meteorológicos, a partir de uma média ponderada.

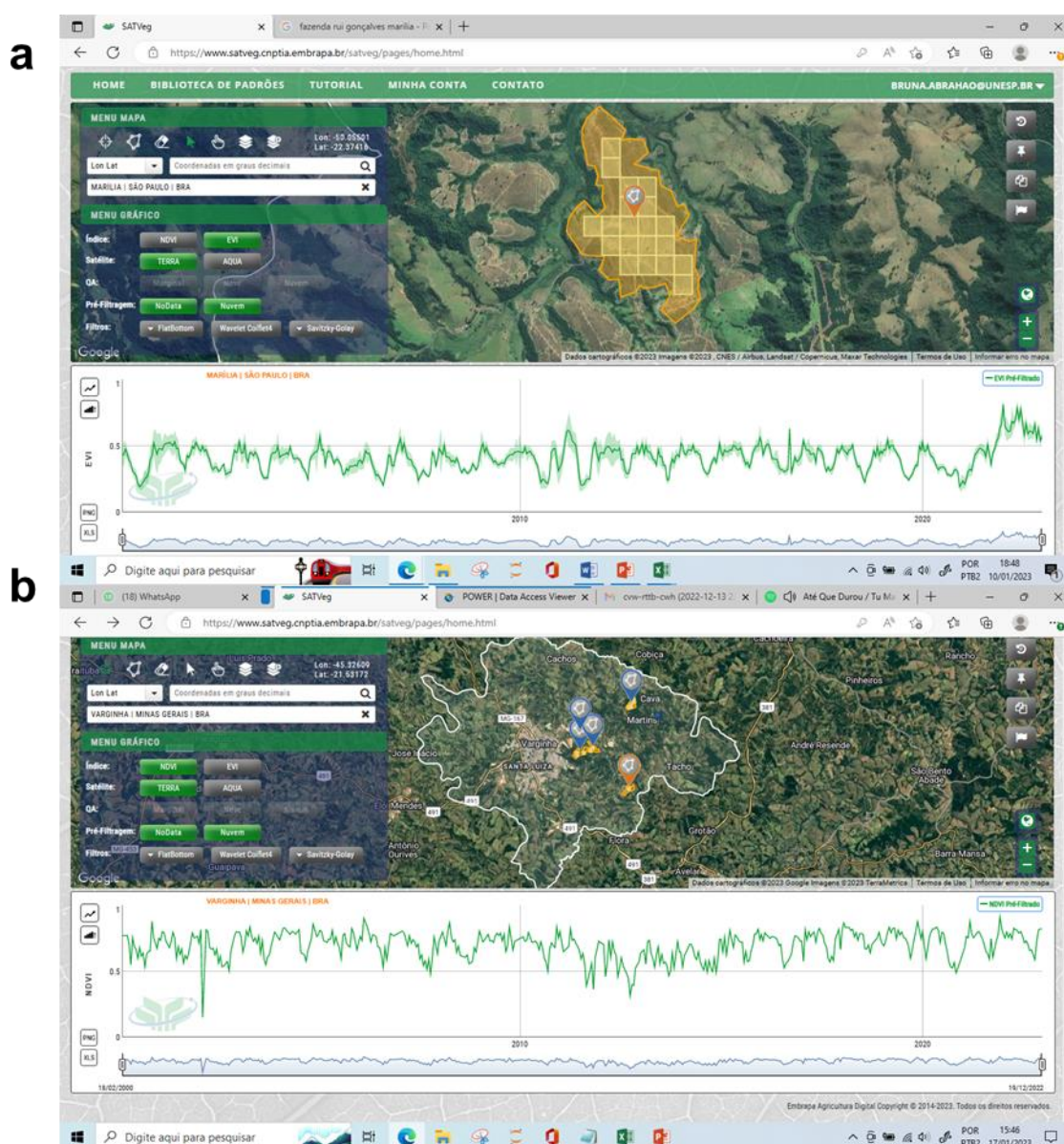


Figura 2- Sistema SATveg da Embrapa. A) Exemplo de coleta de dados em diferentes polígonos, B) Coleta de dados em um município.

2.3. Dados da Cercosporiose

Os dados da severidade da cercosporiose foram coletados a partir do monitoramento mensal dos boletins fitossanitários realizados pela Fundação Procafé (<http://fundacaoprocafe.com.br/>), abrangendo o período 2015 a 2022.

Para a coleta dos dados os cafeeiros foram selecionados seguindo um padrão de ziguezague na área estudada, conforme orientado por Chalfoun (1997). Em cada talhão, foram amostradas 20 plantas, sendo que em cada lado da planta foram selecionados cinco ramos laterais. Após isso, foi removida uma folha completamente desenvolvida, do 3º ou 4º par de folhas, de cada ramo, resultando na coleta de um total de 10 folhas por planta (cinco de cada lado). Com isso, foram analisadas 200 folhas por talhão para determinar a severidade da cercosporiose no café.

Os dados de Severidade (SEV%) mensais, obtidos pela fundação Procafé são classificados de acordo com áreas de alta (carga alta) e baixa (carga baixa) produtividade, considerando as variações nos tratos culturais aplicados na região e da bienalidade da produção. O ano de carga alta, o café concentra sua energia na produção de grãos, portanto tem uma alta produtividade, enquanto em um ano de carga baixa, o café concentra suas energias no crescimento vegetativo, com isso, apresenta baixa produtividade.

2.4. Modelos de Estimativas

A estimativa da severidade da cercosporiose, foi baseada em dados de severidade da doença fornecidos pela Fundação Procafé, em conjunto com os dados meteorológicos extraídos da plataforma Nasa Power e os índices de vegetação (NDVI e EVI) obtidos da plataforma SATveg. Para realizar essa estimativa, foram aplicados três modelos: Árvore de Decisão (AD), Regressão Linear Múltipla (RLM) e Random Forest (RF).

Os dados meteorológicos e os índices de vegetação foram utilizados como variáveis independentes, enquanto os dados da cercosporiose foram tratados como a variável dependente em todos os modelos testados.

A regressão linear múltipla é uma técnica multivariada que tem o objetivos de estabelecer uma relação matemática entre a variável dependente e variáveis independentes que descrevem o sistema. Além disso, pode reduzir muitas variáveis a poucas dimensões, com poucas perdas de informações, o que facilita a identificação dos principais padrões de similaridade, associação e correlação entre as variáveis. Após estabelecida a relação matemática, sua principal aplicação é gerar valores para a variável dependente, com base nas variáveis independentes (SASSI et, al. 2012).

O modelo de Árvore de Decisão (AD) é uma representação gráfica em formato de árvore invertida, composta por nós e ramos. (MEIRA et al., 2008; DELGADO et al., 2012). O objetivo principal de uma árvore de decisão é criar um modelo preditivo preciso ou descobrir a estrutura subjacente do problema estudado (BREIMAN et al., 1984). Uma das principais vantagens das Árvores de Decisão é sua capacidade de considerar os atributos mais relevantes e apresentá-los de maneira clara e compreensível, facilitando a identificação dos fatores mais influentes no estudo (CREPALDI et, al. 2012)

Random Forest (RF) é uma técnica de aprendizado supervisionado, que consiste em um conjunto de árvores de decisão (VIEIRA, 2015). Este algoritmo foi desenvolvido para melhorar o desempenho de preditivo em relação a uma única árvore de decisão (GHOSH et al., 2019; RIBEIRO et al., 2020). O

classificador de Random Forest gera múltiplas árvores de decisão, utilizando subconjuntos aleatórios de amostras e variáveis (KHAIDEM et al., 2016).

O Random Forest apresenta como vantagem ser relativamente rápido para treinar e realizar previsões, além de necessitar de poucos parâmetros de ajuste. O método fornece recursos estatísticos como a importância de cada variável no conjunto de dados, ponderação diferencial de classes, e ferramentas para visualização e detecção de outliers (CUTLER et al., 2011).

Para avaliar a precisão, tendência e acurácia dos modelos de Regressão Linear Múltipla, Árvore de Decisão e Random Forest na estimativa da severidade da cercosporiose, os dados da fase de treinamento serão analisados utilizando métricas estatísticas de desempenho. Serão utilizadas as seguintes métricas: coeficiente de determinação ajustado (R^2 ajustado), erro do quadrado médio (RMSE) e coeficiente de Matthews (MCC), conforme descrito nas Eqs. 4-6.

$$R^2_{ajustado} = 1 - \left[\frac{(1 - R^2)(n - 1)}{n - k - 1} \right] \quad (Eq. 4)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_i^m - Y_i^p)^2} \quad (Eq. 5)$$

$$MCC = \frac{TP * TN - FP * FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (Eq. 6)$$

em que, TP representa o número de verdadeiros positivos, TN representa o número de verdadeiros negativos, FP representa o número de falsos positivos, e FN representa o número de falsos negativos.

3. RESULTADOS

No trabalho foram utilizados três modelos agrometeorológicos, quais foram Regressão Linear Múltipla (RLM), Árvore de decisão e o Random Forest para definir qual seria o melhor modelo para estimar a severidade na cercosporiose no cultivo do café.

Os modelos RLM não tiveram desempenhos satisfatórios para cargas alta e baixa, com valores máximos de R^2 ajustado e RMSE iguais a 0,07 e 9,43, respectivamente, apesar de ser significativo a 0,05 (Figura 3).

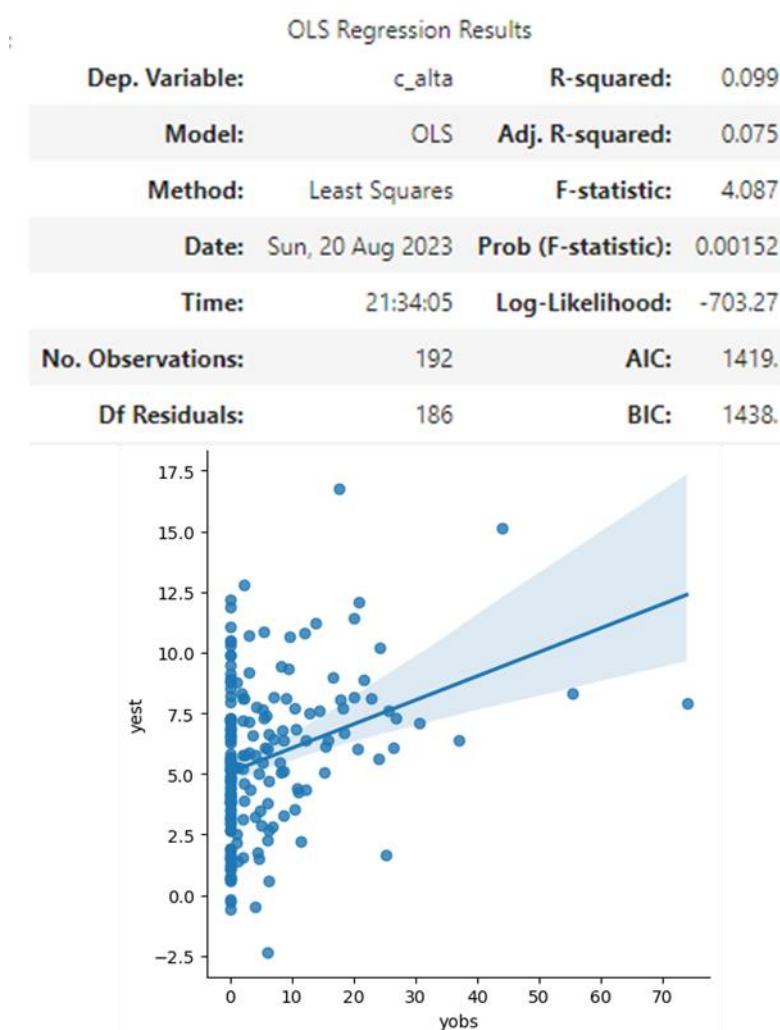


Figura 3. Resultados estatísticos do ajuste do modelo de regressão linear múltiplo para carga alta.

O modelo de RLM resultante foi: $\%SEV = 5,62 - 3,02Qg11 + 3,23TPO6 - 2,53RH2M12 - 0,96Chuva5 + 1,04WS2M11$. Resultados mais completos em relação ao ajuste podem ser verificados na Figura 4.

	coef	std err	t	P> t	[0.025	0.975]
const	5.6211	0.691	8.129	0.000	4.257	6.985
Qg11	-3.0215	0.763	-3.961	0.000	-4.527	-1.516
TPO 6	3.2370	1.176	2.753	0.006	0.917	5.557
RH2M 12	-2.5539	1.038	-2.460	0.015	-4.602	-0.506
Chuva 5	-0.9621	0.759	-1.267	0.207	-2.460	0.536
WS2M 11	1.0456	0.762	1.372	0.172	-0.458	2.549

Figura 4. Resultados do ajuste dos parâmetros da regressão linear múltipla. Em que: const= constante, Qg11 = radiação global de 3 dias antes da avaliação em campo (final do mês), TPO6= temperatura do ponto de orvalho de 8 dias antes da avaliação em campo, RH2M12= umidade relativa do ar de 2 dias antes da avaliação em campo, Chuva5= precipitação de 9 dias antes da avaliação em campo e WS2M1= velocidade do vento de 3 dias antes da avaliação em campo.

Devido ao baixo desempenho de RLM, optou-se por trabalhar com a SEV% separada em classes, que consistem em: Nula ($SEV\% = 0$), Baixa ($0 < SEV\% \leq 8\%$), Média ($8\% < SEV\% < 15\%$) e Alta ($SEV\% > 15\%$). Assim o modelo passa a estimar classes e não valores, aumentando a acurácia dos modelos.

Os modelos de Árvore de Decisão (AD) passaram a ter maior acurácia para estimação da doença tanto para cargas altas (figura 5 e 6) como para cargas baixas (figuras 7 e 8), o que pode ser verificado por índices de Matthew (MCC) sempre superiores a 0,84 e 0,9 para carga alta e carga baixa, respectivamente.

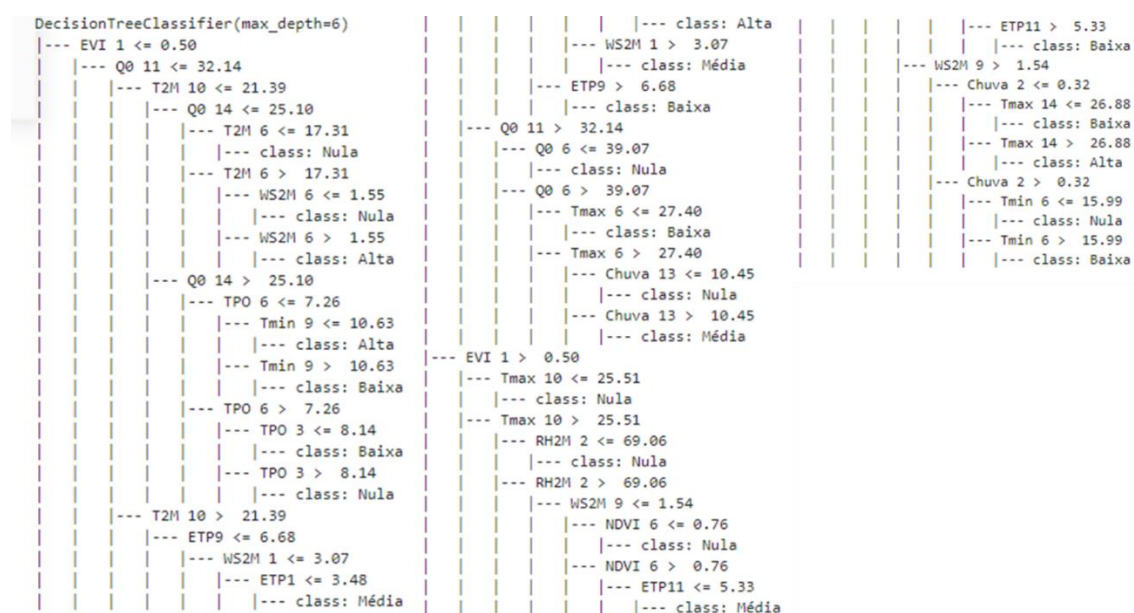


Figura 5. Modelo de Árvore de Decisão ajustado para Carga Alta

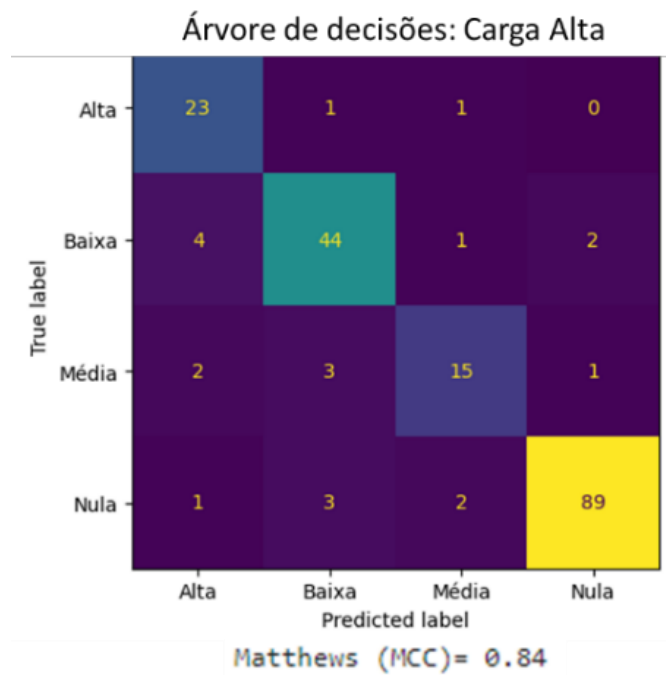


Figura 6. Matriz de Confusão dos resultados do modelo de árvore de decisão para carga alta

O modelo de AD para carga baixa (figura 7) foi acurado pois o coeficiente de Matthews foi igual a 0,91 (figura 8).

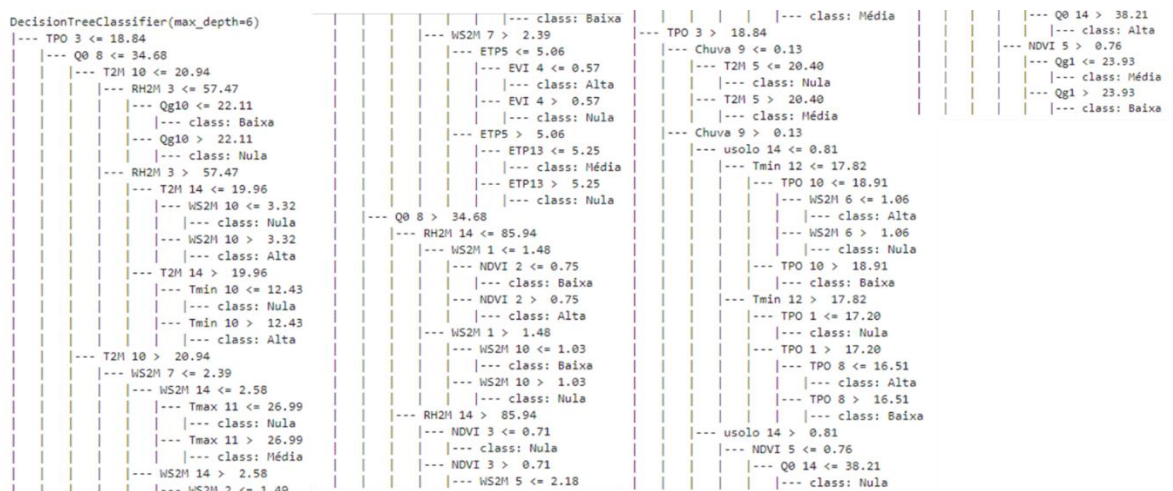


Figura 7. Modelo de Árvore de Decisão ajustado para Carga Baixa

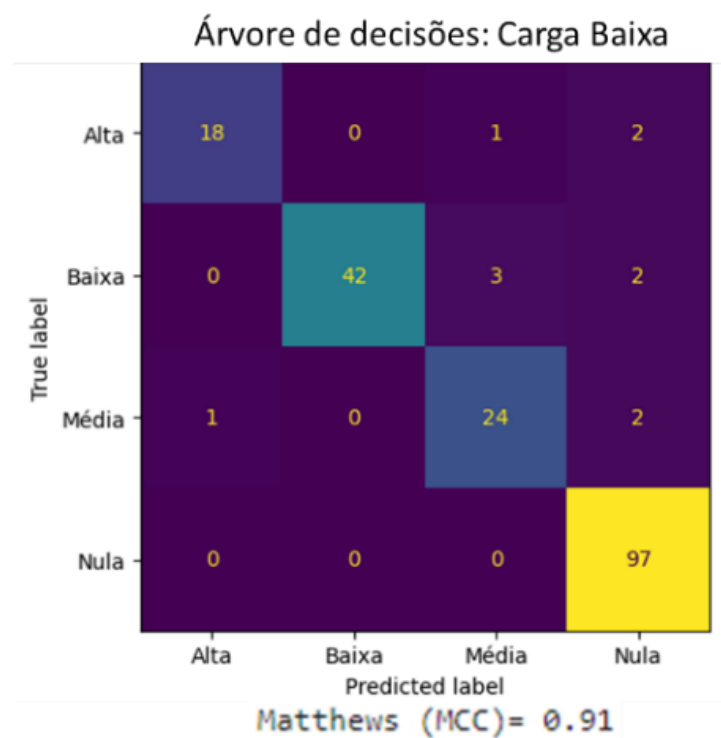


Figura 8. Matriz de Confusão dos resultados do modelo de árvore de decisão para carga baixa

Para o modelo de RF o desempenho foi elevado com $MCC = 0,91$ (figura 9) para carga alta e $MCC = 0,9$ (figura 10) para carga baixa, demonstrando o elevado desempenho do modelo.

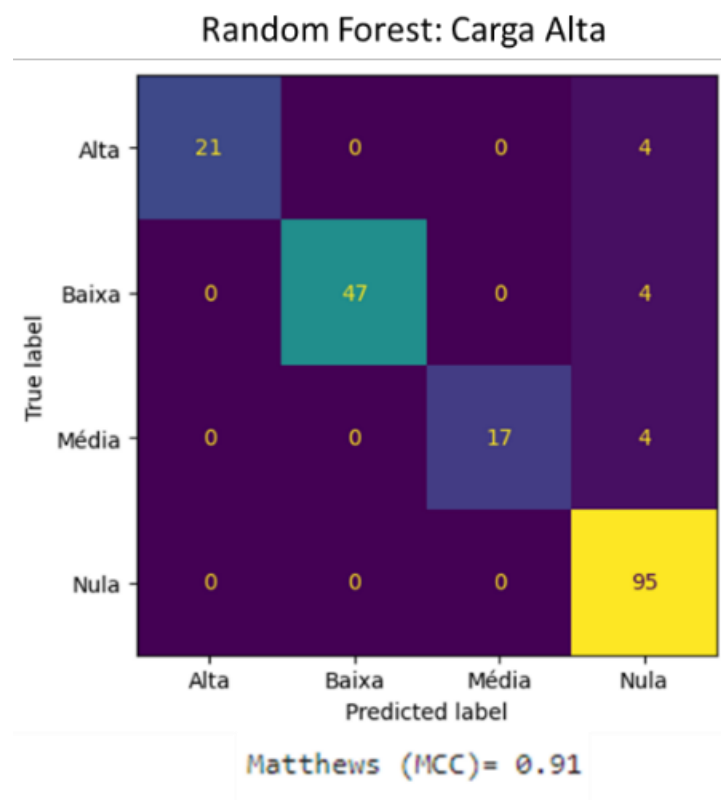


Figura 9. Matriz de Confusão dos resultados do modelo de Random Forest para carga alta

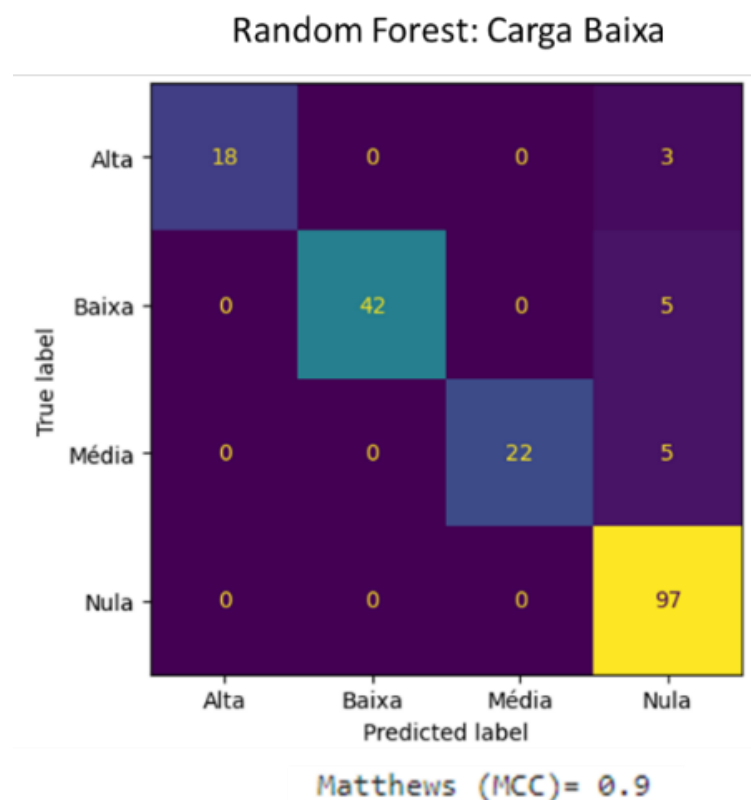


Figura 10. Matriz de Confusão dos resultados do modelo de Random Forest para carga baixa

Assim observou-se que os modelos de machine learning AD e RF tiveram melhores desempenhos que o de RLM permitindo uma previsão de severidade de cercospora. Esses resultados estão de acordo com Carvalho (2020) que obtiveram resultados semelhantes para a produtividade de Erva-Mate na região sul do Brasil. De acordo com o trabalho desenvolvido por Carvalho (2020), o modelo de RLM desenvolvido não foi capaz de capturar a variabilidade dos dados para fazer estimações, assim ela utilizou o modelo de Random Forest para obter modelos mais acurados.

4. CONCLUSÃO

É possível estimar a severidade cercosporiose a partir de variáveis meteorológicas e de sensoriamento remoto. O modelo de random forest teve maior desempenho para estimar a severidade cercosporiose se comparado com a regressão linear múltipla e de árvore de decisão.

5. REFERÊNCIAS

- ACEVES-NAVARRO, L. A.; RIVERA-HERNÁNDEZ, B.; SANTILLÁN-FERNÁNDEZ, A.; ARRIETA-RIVERA, A.; JUÁREZ-LÓPEZ, J. F.; GUTIÉRREZ-BURÓN, R. Impacto del cambio climático en la adaptación del cultivo de café (*Coffea arabica* L.) en Tabasco, México. **Agroproductividad**, v.13, n.4, p.53–58, 2020.
- AGRIANUAL. **Anuário da Agricultura Brasileira**. São Paulo: FNP Consultoria e Comércio, 521 p., 1999.
- ALÈGRE, C. Climats et caféiers d'Arabie. **Agronomie Tropicale**, v.14, p.23-58, 1959.
- ALLEN, R. G.; PEREIRA, L. S.; RAES, D.; SMITH, M. **Crop evapotranspiration: Guidelines for computing crop water requirements**. Rome: FAO, 300 p., 1998. (FAO – Irrigation and Drainage Paper, 56).
- ASSAD, E. D.; PINTO, H. S.; JUNIOR, J. Z.; ÁVILA, A. M. H. Impacto das mudanças climáticas no zoneamento agroclimático do café no Brasil. **Pesquisa Agropecuária Brasileira**, Brasília, v. 39, n. 11, p. 1057- 1064, nov. 2004.
- BERUSKI, G. C.; PEREIRA, A. B.; JACCOUD-FILHO, D. S.; SARTORI, F. F.; SENTELHAS, P. C. Incidence and severity of white mold for soybean under different cultural practices and local meteorological conditions. **Bioscience Journal**, Uberlândia, v. 31, n. 4, p. 1004-1014, 2015.
- BOTELHO, D. M. S.; RESENDE, M. L. V.; ANDRADE, V. T.; PEREIRA, A. A.; PATRICIO, F. R. A.; JUNIOR, P. M. R.; OGOSHI, C.; REZENDE, J. C.

- Cercosporiosis resistance in coffee germplasm collection. **Euphytica**, v. 213, n. 6, p. 117, 2017.
- BREIMAN, L.; FRIEDMAN, J. H.; OLSHEN, R. A.; STONE, C. J. Classification and regression trees. Boca Raton: CRC Press, 358 p., 1984.
- CAMARGO, M. B. P. The impact of climatic variability and climate change on arabic coffee crop in Brazil. **Bragantia**, Campinas, v. 69, p. 239-247, 2010.
- CARVALHO, C. H. S. **Cultivares de café: origem, características e recomendações**. Brasília: Embrapa-Café, 334 p., 2008.
- CARVALHO, M. J. N. **Modelagem Agrometeorológica para estimação de produtividade de erva-mate (*ilex paraguariensis* ST.HIL.) na região sul do Brasil**. 2020. 61f. Dissertação (Mestrado em Agronomia) – Faculdade de Ciências Agrárias e Veterinárias de Jaboticabal, Universidade Estadual Paulista “Júlio de Mesquita Filho”, Jaboticabal, 2020.
- CHALFOUN, N. Design and application of natural down-draft evaporative cooling devices. Boulder: American Solar Energy Society Inc, 1997.
- COMPANHIA NACIONAL DE ABASTECIMENTO [CONAB]. **Acompanhamento da safra brasileira de café: terceiro levantamento**. Setembro 2020. Disponível em: <https://www.conab.gov.br/infoagro/sagra/cafe>. Acesso em: 07/2023
- COMPANHIA NACIONAL DE ABASTECIMENTO [CONAB]. **Acompanhamento da safra brasileira de café: terceiro levantamento**. Janeiro 2021. Disponível em: <https://www.conab.gov.br/infoagro/sagra/cafe>. Acesso em: 07/2023

- CREPALDI, P. G.; AVILA, R. N. P.; OLIVEIRA, J. P. M.; RODRIGUES, P. R.; MARTINS, R. L. **Um estudo sobre a árvore de decisão e sua importância na habilidade de aprendizado**. 2010. Disponível em: https://www.inesul.edu.br/revista/arquivos/arq-idvol_15_1320100263.pdf. Acesso em: 08/2023
- CUTLER, A.; CUTLER, D.; STEVENS, J. **Random Forest**, 2011. Disponível em: <https://www.researchgate.net/publication>. Acesso em: 08/2023
- DAMATTA, F. M.; RENA, A. B. Ecofisiologia de cafezais sombreados e a pleno sol. In: ZAMBOLIM, L. **O Estado da arte de tecnologias na produção de café**. Viçosa: UFV, 2002. p. 93-135.
- DELGADO, R. C. SEDIYAMA, G. C.; COSTA, M. H.; SOARES, V. P.; ANDRADE, R. G. Classificação espectral de área plantada com a cultura da cana-de-açúcar por meio da árvore de decisão. **Engenharia Agrícola**, v. 32, n. 2, p. 369-380, 2012.
- GHOSH, I; JANA, R. K; SANYAL, M. K. Analysis of temporal pattern, causal interaction and predictive modeling of financial markets using nonlinear dynamics, econometric models and machine learning algorithms. **Applied Soft Computing**, 82, 105553, 2019.
- GILLESPIE, T. J.; SENTELHAS, P. C. Agrometeorology and plant disease management: a happy marriage. **Scientia Agricola**, v. 65, n. spe, p. 71-75, 2008.
- GODOY, C. V.; BERGAMIN FILHO, A.; SALGADO, C. L. Doenças do cafeeiro. In: **Manual de fitopatologia: doenças das plantas cultivadas**. São Paulo: Agronômica Ceres, v. 2, p. 184-200, 1997. Cap. 17.

- HARVEY, C. A.; SABORIO-RODRÍGUEZ, M.; MARTINEZ-RODRÍGUEZ, M. R.; VIGUERA, B.; CHAIN-GUADARRAMA, A. Climate change impacts and adaptation among smallholder farmers in Central America. **Agriculture & Food Security**, v. 7, n. 1, p. 1–20, 2018.
- HODGES, T. Predicting Crop Phenology. Boca Raton: **CRC Press**, 1990.
- KATSUHAMA, N.; IMAI, M.; NARUSE, N.; TAKAHASHI, Y. Discrimination of areas infected with coffee leaf rust using a vegetation index. **Remote Sensing Letters**, v. 9, n. 12, p. 1186-1194, 2018.
- KHAIDEM, L; SAHA, S; DEY, S. R. Predicting the direction of stock market prices using random forest. **Applied Mathematical Finance**. 2016
- LI, J. et al. Diversifying crop rotation improves system robustness. **Agronomy for Sustainable Development**, v.39, n.4, p.1–13, 2019.
- LIMA, L. M., POZZA, E. A., SANTOS, F. S. Relationship between incidence of brown eye spot of coffee cherries and the chemical composition of coffee beans. **Journal of Phytopathology**, Berlin, v. 160, n. 4, p. 209-211, 2012
- KUCKENBERG, J.; TARTACHNYK, I.; NOGA, G. Temporal and spatial changes of chlorophyll fluorescence as a basis for early and precise detection of leaf rust and powdery mildew infections in wheat leaves. **Precision agriculture**, v. 10, n. 1, p. 34-44, 2009.
- MALTHUS, T. J.; MADEIRA, A. C. High resolution spectroradiometry: spectral reflectance of field bean leaves infected by *Botrytis fabae*. **Remote Sensing of Environment**, v. 45, n. 1, p. 107-116, 1993
- MARTINS, G. D.; GALO, M. L. B. T.; VIEIRA, B. S. Detecting and mapping root-knot nematode infection in coffee crop using remote sensing

- measurements. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, v. 10, n. 12, p. 5395-5403, 2017.
- MATIELLO, J. B. O café: do cultivo ao consumo. São Paulo: **Globo**, 1991. 320 p. (Coleção do Agricultor - Grãos).
- MATIELLO, J. B.; SANTINATO, R.; GARCIA, A. W. R.; ALMEIDA, S. R.; FERNANDES, D. R. **Cultura de café no Brasil: manual de recomendações**. Ed. 1 Varginha: Fundação Procafé, 2010.
- MEIRA, C. A. A.; RODRIGUES, L. H.; MORAES, S. A. Análise da epidemia da ferrugem do cafeeiro com árvore de decisão. **Tropical Plant Pathology**, v. 33, p. 114-124, 2008.
- MIRIK, M.; MICHELS JR., G. J.; KASSYMZHANOVA-MIRIK, S.; ELLIOTT, N. C.; CATANA, V.; JONES, D. B.; BOWLING, R. Using digital image analysis and spectral reflectance data to quantify damage by greenbug (Hemiptera: Aphididae) in winter wheat. **Computers and electronics in agriculture**, 51(1-2): 86-98, 2006
- NASA POWER. **Meteorological**. Washington, D.C.: National Aeronautics and Space Administration, 2021. Disponível em: [https://power.larc.nasa.gov/docs/methodology /data/sources/](https://power.larc.nasa.gov/docs/methodology/data/sources/). Acesso em: 7 ago. 2023.
- POZZA, A. A. A., PRIETO MARTINEZ, H. E., CAIXETA, S. L., CARDOSO, A. A., ZAMBOLIM, L., & POZZA, E. A. Influência da nutrição mineral na intensidade da mancha-de-olho-pardo em mudas de cafeeiro. **Pesquisa Agropecuária Brasileira**, Brasília, v. 36, n. 1, p. 53-60, 2001.

- POZZA, E. A., ALVES, M. C. Impacto potencial de mudanças climáticas sobre as doenças fúngicas do cafeeiro no Brasil. In: Ghini R, Ramada E. (Ed.). **Mudanças climáticas: impactos sobre doenças de plantas no Brasil**. Brasília: EMBRAPA, 220-238, 2008.
- RIBEIRO, M. H. D. M; COELHO, L. S. Ensemble approach based on bagging, boosting and stacking for short-term prediction in agribusiness time series. **Applied Soft Computing**, 86, 105837, 2020.
- ROLIM, G. S., DE OLIVEIRA APARECIDO, L. E., DE SOUZA, P. S., LAMPARELLI, R. A.C., & DOS SANTOS, É. R. Clima e qualidade natural de Coffea arabica L. bebida. **Theoretical and Applied Climatology**, 1-12, 2020.
- ROLIM, G. S., RIBEIRO, R. V., AZEVEDO, F. A., CAMARGO, M. B. P., MACHADO, E. Predicion of fruits number from the quantity of reproductive structures of orange. **Rev Bras Frutic** 30(1):48–53, 2008.
- SAHOO, B. B. et al. Long short-term memory (LSTM) recurrent neural network for low-flow hydrological time series forecasting. **Acta Geophysica**, v.67, n.5, p.1471–1481, 2019.
- SALGADO, B. G.; MACEDO, R. L. G.; CARVALHO, V. L.; SALGADO, M.; VENTURIN, N. Progresso da ferrugem e da cercosporiose do cafeeiro consorciado com grevilea, com ingazeiro e a pleno sol em Lavras-MG. **Ciência e agrotecnologia**, v. 31, p. 1067-1074, 2007.
- SASSI, C. P.; PEREZ, F. G.; MIYAZATO, L.; XIAO, Y.; SILVA, P. H. F.; LOUZADA, F. Modelos de regressão linear múltipla utilizando os

- softwares R e statistica: uma aplicação a dados de conservação de frutas.
São Carlos: ICMC-USP, 2012.
- SENTELHAS, P. C.; FARIA, R. T.; CHAVES, M. O.; HOOGENBOOM, G.
Evaluation of the WGEN and SIMMETEO weather generators for the
brazilian tropics and subtropics, using crop simulation models. **Revista
Brasileira de Agrometeorologia**, v. 9, n. 2, p. 357-376, 2001.
- SENTELHAS, P. C.; MARTA, A. D.; ORLANDINI, S.; SANTOS, E. A.;
GILLESPIE, T. J.; GLEASON, M. L. Suitability of relative humidity as an
estimator of leaf wetness duration. **Agricultural and forest meteorology**,
v. 148, n. 3, p. 392-400, 2008.
- SILVA, H. R.; POZZA, E. A.; SOUZA, P. E.; FERREIRA, M. A.; FREITAS, A. S.;
MOREIRA, S. I. Cercospora leaf spot in *Toona ciliata*: Epidemiology and
infection process of *Cercospora* cf. *alchemillicola*. **Forest Pathology**,
v.48, n.6, e12451, 2018.
- SOUZA, A. G. C.; MAFFIA, L. A.; MIZUBUTI, E. S. G. Cultural and
Aggressiveness Variability of *Cercospora coffeicola*. **Journal of
Phytopathology**, v.160, n.10, p.540–546, 2012.
- THOMAZIELLO, R. A., FAZUOLI, L. C., PEZZOPANE, J. R., FAHL, J. I.,
CARELLI, M. L. C. **Café arabica: cultura e técnicas de produção**.
Campinas: 2000. Instituto Agrônômico, 82 p. (Boletim Técnico n.º 187)
- THORNTHWAITE, C. W.; MATHER, J. R. The water balance. centerton: Drexel
institute of technology. **Climatology**, 8(1), 104p, 1955
- VIEIRA, G. **Utilização de classificador random forest na detecção de falhas
em máquinas rotativas**. 2015. 79f. Trabalho de Conclusão de Curso

(Graduação) – Curso de Engenharia Eletrônica e de Computação da Escola Politécnica, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2015

WEST, J. S.; BRAVO, C.; OBERTI, R.; LEMAIRE, D.; MOSHOU, D.; McCARTNEY, H. A. The potential of optical canopy measurement for targeted control of field crop diseases. **Annual review of Phytopathology**, v. 41, n. 1, p. 593-614, 2003

XU, X. Effects of environmental conditions on the development of Fusarium ear blight. Em: XU, X.; BAILEY, J. A.; COOKE, B. M. (Eds.). **Epidemiology of Mycotoxin Producing Fungi**. Dordrecht: Springer Netherlands, 2003. p.683–689.

ZAMBOLIM, L.; VALE, F. X. R.; ZAMBOLIM, E. M. Doenças do cafeeiro. In: KIMATI, H. et al. (Ed.). **Manual de fitopatologia: doenças de plantas cultivadas**. 4. ed. São Paulo: Ceres, v. 2, p. 165- 180, 2005

ZELBER-SAGI, S.; SALAMONE, F.; WEBB, M.; LOTAN, R.; YESHUA, H.; HALPERN, Z.; SANTO, E.; OREN, R.; SHIBOLET, O. Coffee consumption and nonalcoholic fatty liver onset: a prospective study in the general population. **Translational Research**, v. 165, n. 3, p. 428-436, 2015.