UNIVERSIDADE ESTADUAL PAULISTA – UNESP

CAMPUS DE JABOTICABAL

MODELOS AGROMETEOROLÓGICOS PARA PREVISÃO DE PRAGAS E DOENÇAS EM Coffea arabica L. EM MINAS GERAIS

Lucas Eduardo de Oliveira Aparecido

Engenheiro Agrônomo

2019

UNIVERSIDADE ESTADUAL PAULISTA – UNESP

CAMPUS DE JABOTICABAL

MODELOS AGROMETEOROLÓGICOS PARA PREVISÃO DE PRAGAS E DOENÇAS EM *Coffea arabica* L. EM MINAS GERAIS

Lucas Eduardo de Oliveira Aparecido

Orientador: Prof. Dr. Glauco de Souza Rolim

Tese apresentada à Faculdade de Ciências Agrárias e Veterinárias – Unesp, Câmpus de Jaboticabal, como parte das exigências para a obtenção do título de Doutor em Agronomia (Produção Vegetal)

2019

A639m

Aparecido, Lucas Eduardo de Oliveira
Modelos agrometeorológicos para previsão de pragas e doenças em Coffea arabica L. em Minas Gerais / Lucas Eduardo de Oliveira Aparecido. -- Jaboticabal, 2019
157 p. : il., tabs., fotos, mapas

Tese (doutorado) - Universidade Estadual Paulista (Unesp), Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal
Orientador: Glauco de Souza Rolim

1. Climatologia agrícola. 2. Big data. 3. Algoritmos. 4. Cafeeiro. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Jaboticabal

CERTIFICADO DE APROVAÇÃO

TÍTULO DA TESE: MODELOS AGROMETEOROLÓGICOS PARA PREVISÃO DE PRAGAS E DOENÇAS EM *Coffea arábica* L. EM MINAS GERAIS

AUTOR: LUCAS EDUARDO DE OLIVEIRA APARECIDO
ORIENTADOR: GLAUCO DE SOUZA ROLIM

Aprovado como parte das exigências para obtenção do Título de Doutor em AGRONOMIA (PRODUÇÃO VEGETAL), pela Comissão Examinadora:

Prof. Dr. GLAUCO DE SOUZA ROLIM
Departamento de Ciências Exatas / FCAV / UNESP - Jaboticabal

Prof. Dr. PAULO SERGIO DE SOUZA
IFSuldeMinas / Muzambinho/MG

Prof. Dr. ALAN RODRIGO PANOSSO
Departamento de Ciências Exatas / FCAV / UNESP - Jaboticabal

Prof. Dr. CICERO TEIXEIRA SILVA COSTA
Laboratório de Irrigação-IFMS-Campus de Naviraí / Naviraí/MS

Prof. Dr. NEWTON LA SCALA JUNIOR
Departamento de Ciências Exatas / FCAV / UNESP - Jaboticabal

Jaboticabal, 30 de agosto de 2019

DADOS CURRICULARES DO AUTOR

LUCAS EDUARDO DE OLIVEIRA APARECIDO – Nascido em 08 de junho de 1992, no município de Nova Resende, Estado de Minas Gerais, Brasil. Ingressou no curso Técnico em Agropecuária no Instituto Federal de Educação Ciência e Tecnologia do Sul de Minas – Campus Muzambinho, em fevereiro de 2007. No mesmo Campus, em fevereiro de 2010, ingressou no curso de Engenharia Agronômica. Durante a graduação participou do Grupo de Pesquisa em Fruticultura, no período de 2010 a 2013. Foi bolsista da Fundação de Amparo a Pesquisa do Estado de Minas Gerais (FAPEMIG), na modalidade de iniciação científica no período de 2011 a 2012. No ano de 2012, estagiou na Empresa de Pesquisa Agropecuária de Minas Gerais (EPAMIG) – Campo Experimental de Caldas e na Universidade Estadual Paulista Júlio Mesquita Filho, no Campus de Jaboticabal, no ano de 2013. Em fevereiro de 2014, obteve o título de Engenheiro Agrônomo. Iniciou o curso de pós-graduação *Stricto sensu* na modalidade Mestrado em Agronomia (Produção Vegetal) em março de 2014, na Universidade Estadual Paulista "Júlio Mesquita Filho", no Campus de Jaboticabal, no Departamento de Ciências Exatas sob a orientação do Prof. Dr. Glauco de Souza Rolim (*GROUP OF AGROMETEOROLOGICAL STUDIES "GAS"*), atuando em pesquisas na área de Agrometeorologia e Modelagem. Em fevereiro de 2016, obteve o título de Mestre em Agronomia (Produção Vegetal). Iniciou o curso de pós-graduação *Stricto sensu* na modalidade Doutorado em Agronomia (Produção Vegetal) em março de 2016 também sob a orientação do Prof. Dr. Rolim. Em junho de 2017, ingressou no serviço público federal no cargo de professor EBTT no Instituto Federal de Mato Grosso do Sul (IFMS-CAMPUS NAVIRAÍ), na área de Engenharia Agrícola, onde leciona até o momento.

“O importante é não parar de questionar. A curiosidade tem a sua própria razão para existir.”

Einstein, Albert

DEDICO

A Deus, o grande doutor da vida.

A minha querida esposa, Adriana Ferreira de Moraes Oliveira, companheira para todos os momentos.

AGRADECIMENTOS

À Universidade Estadual Paulista (UNESP) – Faculdade de Ciências Agrárias e Veterinárias, Campus de Jaboticabal, pela oportunidade concedida para a realização do Mestrado e Doutorado em Agronomia (Produção Vegetal).

Ao Programa de Agronomia (Produção Vegetal) e a todos os professores que, ao longo da minha formação, tive o privilégio de conviver e que foram decisivos para que eu chegasse até este momento.

Ao professor Dr. Glauco de Souza Rolim, pela orientação, confiança, incentivo e parceria ao longo dos anos de pós-graduação. Sua orientação durante esta caminhada e todas as demais, sem dúvida, foi determinante para minha formação.

Ao professor Dr. Paulo Sergio de Souza, pela orientação e parceria desde a época do curso Técnico em Agropecuária. Sua orientação inicial foi determinante para minha formação de hoje.

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pela concessão de recurso financeiro e bolsa para a realização desta pesquisa, Processo nº 2015/17797-4.

Ao Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso do Sul (IFMS), Campus de Naviraí, na qual iniciei e desempenho minhas atividades de docência.

A minha esposa Adriana Ferreira de Moraes Oliveira, acima de tudo uma grande companheira.

Aos grandes amigos, prof. Dr. Paulo Sergio de Souza e Gentil Luiz Miguel Filho, por todo apoio e ensinamento durante a realização do curso técnico em agropecuária, graduação e pós-graduação.

Aos meus pais, José Antônio Aparecido e Maria Regina de Oliveira Aparecido, e irmão, Renan Gabriel de Oliveira Aparecido pelo apoio, confiança investidos em minha formação.

Aos pais de minha esposa, Lourenço Afonso de Moraes e Tereza Galdina de Moraes, que sempre me ajudaram e apoiaram na minha formação pessoal e profissional.

Aos amigos de Naviraí-MS, José Reinaldo da Silva Cabral de Moraes, Guilherme Botega Torsoni, Cícero Teixeira Silva Costa, Alisson Gaspar Chiquitto, grandes amigos, que me apoiaram durante todos esses anos de docência.

Aos amigos de Jaboticabal-SP, Alexson Filgueiras Dutra, Victor Brunini Moreto, Kamila Cunha de Meneses, Taynara Valeriano, grandes amigos, que me apoiaram durante os anos de pós-graduação.

Ao Departamento de Ciências Exatas, principalmente a Zezé, Shirley e Adriana, aos professores e colegas que de alguma forma ajudaram a enxergar de forma diferente algum problema no decorrer do curso.

Aos meus queridos discentes, do curso Técnico em Agricultura e também da Engenharia Agronômica do IFMS, que sempre dedico e busco melhorar a forma e a maneira de ensiná-los.

Enfim, agradeço a todos, professores, amigos e conhecidos que em algum momento da vida contribui para minha formação, tanto no curso técnico, como na graduação e na elaboração deste trabalho.

Muito Obrigado!

SUMÁRIO

MODELOS AGROMETEOROLÓGICOS PARA PREVISÃO DE PRAGAS E DOENÇAS EM *Coffea arabica* L. EM MINAS GERAIS

RESUMO: O café é a bebida mais consumida no mundo e uma das principais causas para a redução da produtividade e qualidade são os problemas fitossanitários. A estratégia mais comum de controle dessas doenças e pragas é a aplicação de fungicidas e inseticidas foliares, dependendo da intensidade dos mesmos na região. Esse método tradicional pode ser melhorado utilizando de sistemas de alertas por meio de modelos de estimativas dos índices de doenças e pragas. Este trabalho tem como OBJETIVOS: A) Calibrar as variáveis meteorológicas: temperatura do ar e precipitação pluviométrica do sistema ECMWF em relação aos dados de reais de superfície mensurados pelo sistema nacional de meteorologia (INMET) para o estado de Minas Gerais; B) Avaliar quais os elementos meteorológicos exercem maior influência nas principais pragas (broca e bicho-mineiro) e doenças (ferrugem e cercosporiose) do cafeeiro arábica nas principais localidades cafeeiras do Sul de Minas Gerais e do Cerrado Mineiro; C) Desenvolver modelos agrometeorológicos para previsão de pragas e doenças em função das variáveis meteorológicas usando algoritmos de machine learning e procurando uma antecipação temporal suficiente para tomada de decisões. MATERIAL E MÉTODOS: Para o objetivo "A" foram utilizados dados climáticos mensais de temperatura do ar (T, ºC) e precipitação pluviométrica (P, mm) provenientes do ECMWF e do INMET no período de 1979 a 2017. A evapotranspiração potencial foi estimada por Thornthwaite (1948) e balanço hídrico por Thornthwaite e Mather (1955). As comparações entre o ECMWF e INMET foram realizadas pelos índices: acurácia (mean absolute percentage error, MAPE, e root mean squared error, RMSE) e precisão (coeficiente de determinação ajustado, R^2adj). Para o objetivo "B" foram utilizados dados climáticos e fitossanitários de Boa Esperança, Carmo de Minas, Muzambinho e Varginha, situadas na região Sul de Minas (SO_{MG}) e as localidades de Araxá, Araguari e Patrocínio situadas na região do Cerrado Mineiro (CE_{MG}). Foram simulados a tendência de progresso das doenças e pragas ao longo de tempo usando modelos não lineares em função do índice térmico acumulado. Também foi estimada dos níveis de infestação de pragas e severidade de doenças usando regressão linear múltipla. A variável dependente foi os níveis de doenças e pragas e as variáveis independentes: graus dias (DD) acumulado, enfolhamento do café estimado por DD e número de nós estimado por DD. Para o objetivo "C" foram utilizados dados climáticos e fitossanitários da SO_{MG} e CE_{MG}. Os algoritmos calibrados e testados para a previsão das doenças e pragas do café foram 1) Regressão linear múltipla, 2) K-Neighbors, 3) Random Forest e 4) Redes Neurais. RESULTADOS E DISCUSSÃO: Os maiores desvios entre P_{INMET} e P_{ECMWF} foram de 75 mm mo^{-1} e ocorreram no verão. O cafeeiro implantado no CE_{MG} tem maiores índices de doenças e pragas em relação ao café do SO_{MG}. O algoritmo random forest foi mais acurado na previsão da ferrugem, cercospora, bicho-mineiro e broca-do-cafeeiro em ambas as regiões. CONCLUSÃO: As variáveis climáticas oriundas do ECMWF são acuradas e podem modelar o balanço hídrico climatológico. É possível simular a tendência e ainda prever os índices de pragas e doenças do café usando como variáveis regressoras os dados climáticos e metodologia o *machine learning*.

PALAVRAS-CHAVE: fitopatologia, modelagem, aprendizado máquina, bigdata.

AGROMETEOROLOGICAL MODELS FOR FORECASTING PESTS AND DISEASES IN *Coffea arabica* L. IN THE STATE OF MINAS GERAIS

ABSTRACT: Coffee is the most consumed beverage in the world, but phytosanitary problems are amongst the main causes of reduced productivity and quality. The application of foliar fungicides and insecticides is the most common strategy for controlling these diseases and pests, depending on their intensity in a region. This traditional method can be improved by using alert systems with models of disease and pest indices. This work has as OBJECTIVES: A) To calibrate the meteorological variables: air temperature and rainfall of the European Center for Medium Range Weather Forecast (ECMWF) in relation to the real surface data measured by the national meteorological system (INMET) for the state of Minas Gerais; B) To evaluate which meteorological elements, and at what time, have a greater influence on the main pests (coffee borer and coffee miner) and diseases (coffee rust and cercosporiosis) of Coffee arabica in the main coffee regions of the South of Minas Gerais and Cerrado Mineiro; C) To develop agrometeorological models for pest and disease prediction in function of the meteorological variables of the South of Minas Gerais and Cerrado Mineiro using algorithms of machine learning with sufficient temporal anticipation for decision making. MATERIAL AND METHODS: To achieve goal "A" we used monthly climatic data (T, ºC) and rainfall (P, mm) from the ECMWF and INMET from 1979 to 2015. Potential evapotranspiration was estimated by Thornthwaite (1948) and water balance by Thornthwaite and Mather (1955). The comparisons between the ECMWF and INMET were performed by the indexes: mean absolute percentage error (MAPE) and precision mean (R2adj). To achieve the goal "B" we use climatic and phytosanitary data of Boa esperança, Carmo de Minas, Muzambinho and Varginha, located in the South of Mines (SO_{MG}) and the Araxá, Araguari and Patrocínio located in the Cerrado Mineiro region (CE_{MG}). We simulate the trend of disease and pest progression over time using nonlinear models as a function of the accumulated thermal index. And we estimated levels of pest infestation and disease severity using multiple linear regression. The dependent variable was the levels of diseases and pests and the independent variables: cumulative days (DD), coffee leafage estimated by DD and number of nodes estimated by DD. To achieve the "C" objective we use the climatic and phytosanitary data of SO_{MG} and CE_{MG}. The algorithms calibrated and tested for the prediction of coffee pests and diseases were: 1) Multiple linear regression, 2) K-Neighbors Regressor, 3) Random Forest Regressor and 4) Artificial Neural Networks. The best models were selected using the MAPE, Willmott's 'd', RMSE and R^2 adj. RESULTS AND DISCUSSION: The largest deviations between P_{INMET} and P_{ECMWF} were 75 mm mo^{-1} and occurred in the summer. The coffee plant implanted in CE_{MG} has higher rates of diseases and pests in relation to SO_{MG} coffee. The random forest algorithm was more accurate in the prediction of coffee rust, cercospora, coffee miner and coffee borer in both regions. CONCLUSION: The climatic variables from the ECMWF are accurate and can be used in modeling the climatological water balance. It is possible to simulate the trend and to predict coffee pests and diseases using as regressive variables the climatic data and machine learning methodology.

KEY- WORDS: plant pathology, modelling, machine learning, bigdata.

CAPÍTULO 1 - Considerações gerais

1 INTRODUÇÃO

1.1 O cultivo do Café no Brasil

O cafeeiro é uma planta perene, pertencente à família Rubiaceae (CUBRY et al., 2013). As duas espécies economicamente mais importantes do café são o *Coffea arabica* L. e *Coffea canefora* Pierre (BRAVO-MONROY et al., 2016), representando 74,92% e 25,08% da produção mundial, respectivamente (CONAB, 2019).

O café é a bebida mais consumida no mundo, apresentando várias propriedades funcionais, como por exemplo, a cafeína, aminoácidos, açúcares e compostos fenólicos (BUTT; SULTAN, 2011; KITZBERGER et al., 2013). Na atividade agrícola brasileira a cafeicultura tem grande importância (RESENDE et al., 2009; RODRIGUES et al., 2013). As áreas de produção brasileira do Café arábica se distribuem na região centro-sul (Figura 1), principalmente nos estados de Minas Gerais, São Paulo, Paraná e Espírito Santo (ANDRADE et al., 2012; CUBRY et al., 2013). Minas Gerais apresenta em torno de 6,9% da área total do Brasil e se destaca como maior produtor de café do país (BARBOSA et al., 2012; RONCHI et al., 2015).

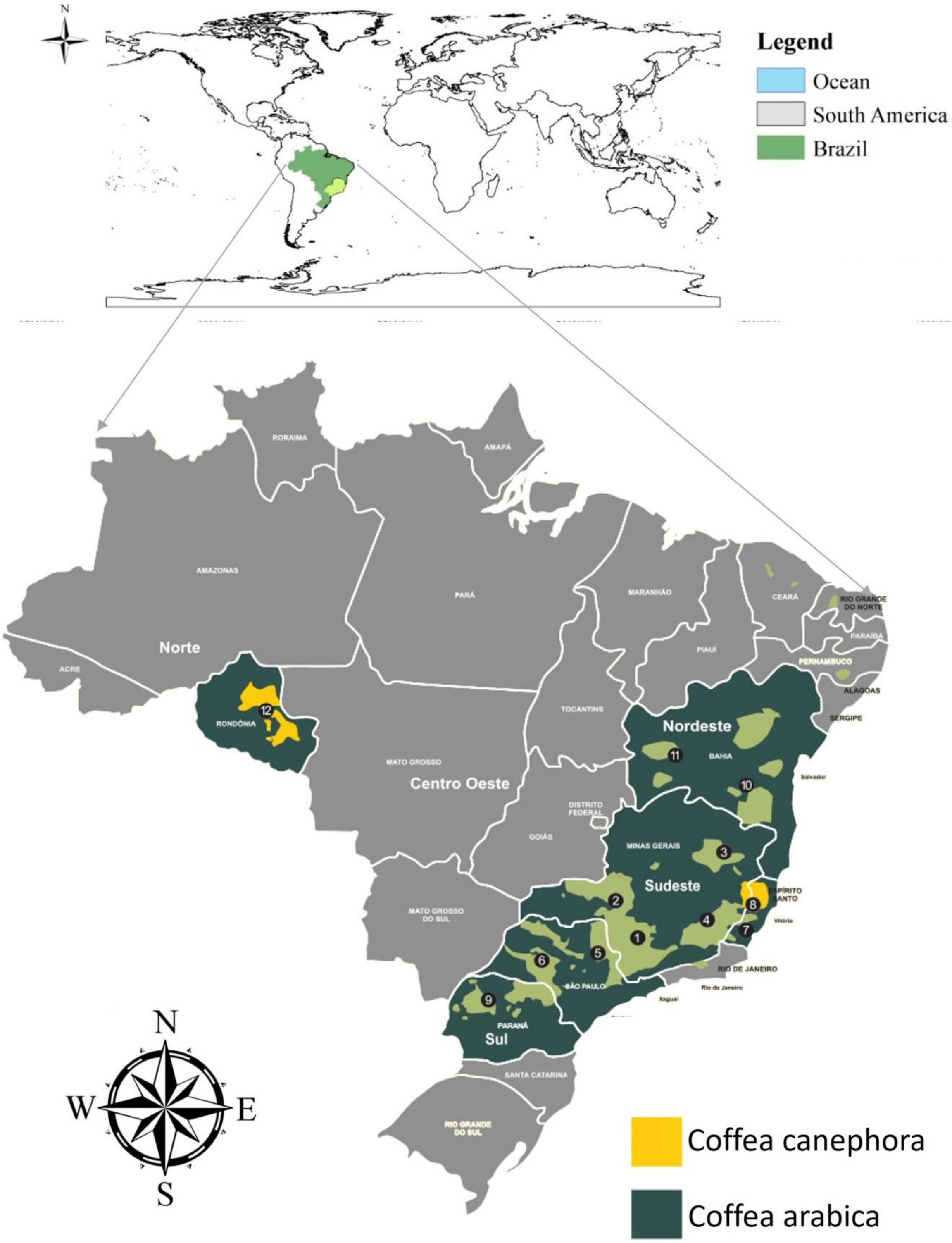

Figura 1. Principais regiões produtoras de café do Brasil (adaptado de ROSSIGNOLLI, 2019; CONAB, 2019.

O cafeeiro arábica é fortemente afetado nos seus diversos estádios fenológicos pelas condições meteorológicas (PICINI et al., 1999) e também pelas doenças e pragas. O cafeeiro arábico é um cultivo que necessita de dois anos para completar todo o seu ciclo fenológico (Figura 2). O período vegetativo ocorre no primeiro ano e o processo de reprodução no segundo ano. No desenvolvimento vegetativo ocorre à formação e o crescimento dos ramos vegetativos, esse processo ocorre em fotoperíodo de dias longos. Com a redução do fotoperíodo as gemas vegetativas axilares são induzidas por fotoperiodismo em gemas reprodutivas. O período produtivo inicia-se com a florada que precede a formação dos chumbinhos. Em agosto inicia-se a expansão e a formação dos grãos que ocorre até atingir o tamanho normal, posteriormente ocorre a granação e a maturação dos frutos (CAMARGO; CAMARGO, 2001).

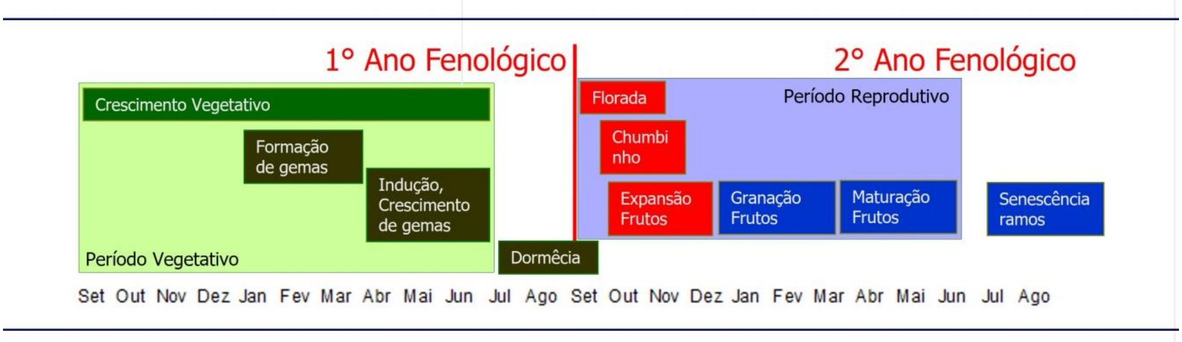

Figure 2. Fenologia do *Coffea arabica* L.

1.2 O clima e as enfermidades do café

A variabilidade climática causa forte impacto nas atividades agrícolas (SÁ JUNIOR et al., 2012), sendo um dos fatores responsáveis pelas flutuações e oscilações das doenças do cultivo, principalmente as doenças de origem fúngicas. Hoogenboom (2000) salienta que os elementos meteorológicos críticos para a produção agrícola são a radiação solar e a precipitação, sendo que, para o café, Camargo (2010) relata que a chuva é o elemento que proporciona maior interferência na fenologia do cafeeiro, e consequentemente, condiciona a intensidade e a severidade das doenças e pragas. Segundo Sentelhas et al., (2007) a temperatura do ar e o molhamento foliar são parâmetros microclimáticos importantíssimos que influenciam a maioria das doenças fúngicas de plantas. A irradiância solar, além de fornecer a energia para a fotossíntese e partição de carboidratos (OLIVEIRA et al., 2012), pode estimular ou inibir o desenvolvimento das doenças do cafeeiro.

Temperaturas do ar entre 5 ºC e 30 ºC com altos níveis de umidade relativa do ar e vapor de água alta na superfície das plantas produzem um molhamento foliar de 42 a 72 horas, o que favorece o desenvolvimento de doenças fúngicas (SENTELHAS et al., 2008; BERUSKI et al., 2015). A duração do molhamento foliar é um parâmetro-chave que influencia a epidemiologia das plantas, pois fornece a água livre exigido por patógenos para infectar o tecido foliar das culturas agrícolas (SENTELHAS et al., 2007).

A temperatura do ar regula a taxa de desenvolvimento dos cultivos e também o nível de esporulação de diversos fungos. Microclimas com temperaturas do ar elevadas juntamente com altos níveis de precipitação durante a floração do cafeeiro proporciona o desenvolvimento de fungos, que levam ao abortamento das flores e consequentemente queda na produção (CAMARGO; PEREIRA, 1994; PEREIRA et al., 2008).

A umidade relativa, a temperatura do ponto de orvalho e o déficit de pressão de vapor interferem na presença e atividade de pragas e doenças (SENTELHAS et al., 2005). A velocidade do vento afeta a taxa transpiratória das plantas e disseminação de insetos e doenças (HOOGENBOOM, 2000). Por isso, monitorar as informações climáticas é tão importante.

1.3 Modelos climáticos globais

O território brasileiro não tem uma rede de estações meteorológicas de superfície que atenda todas as necessidades agrícolas. Assim, na falta de dados meteorológicos é recomendado a utilização de “Global Circulation Models”. São exemplos desses modelos os dados do European Centre for Medium Range Weather Forecasts (ECMWF) e do Prediction of Worldwide Energy Resources

(NASA POWER).

Os modelos que fazem previsões das condições de tempo e clima são ferramentas importantes para auxiliar nas modelagens de cultivos, como no caso de previsão de produtividade ou monitoramento do cultivo (BECHTOLD et al., 2008). Os dados do ECMWF, principalmente o ERA-interim, são muito utilizado para previsões de tempo e clima no mundo, sendo disponibilizado dados de previsão em um grid de 0,25 ° (aproximadamente 25 x 25 km) para o mundo todo, tornando-se assim interessante para obtenção de dados em regiões com escassez de estações meteorológicas de superfície (MORAES et al., 2012). Com o ERA-interim é possível contemplar todo território brasileiro com uma malha de 11.358 estações virtuais (Figura 3).

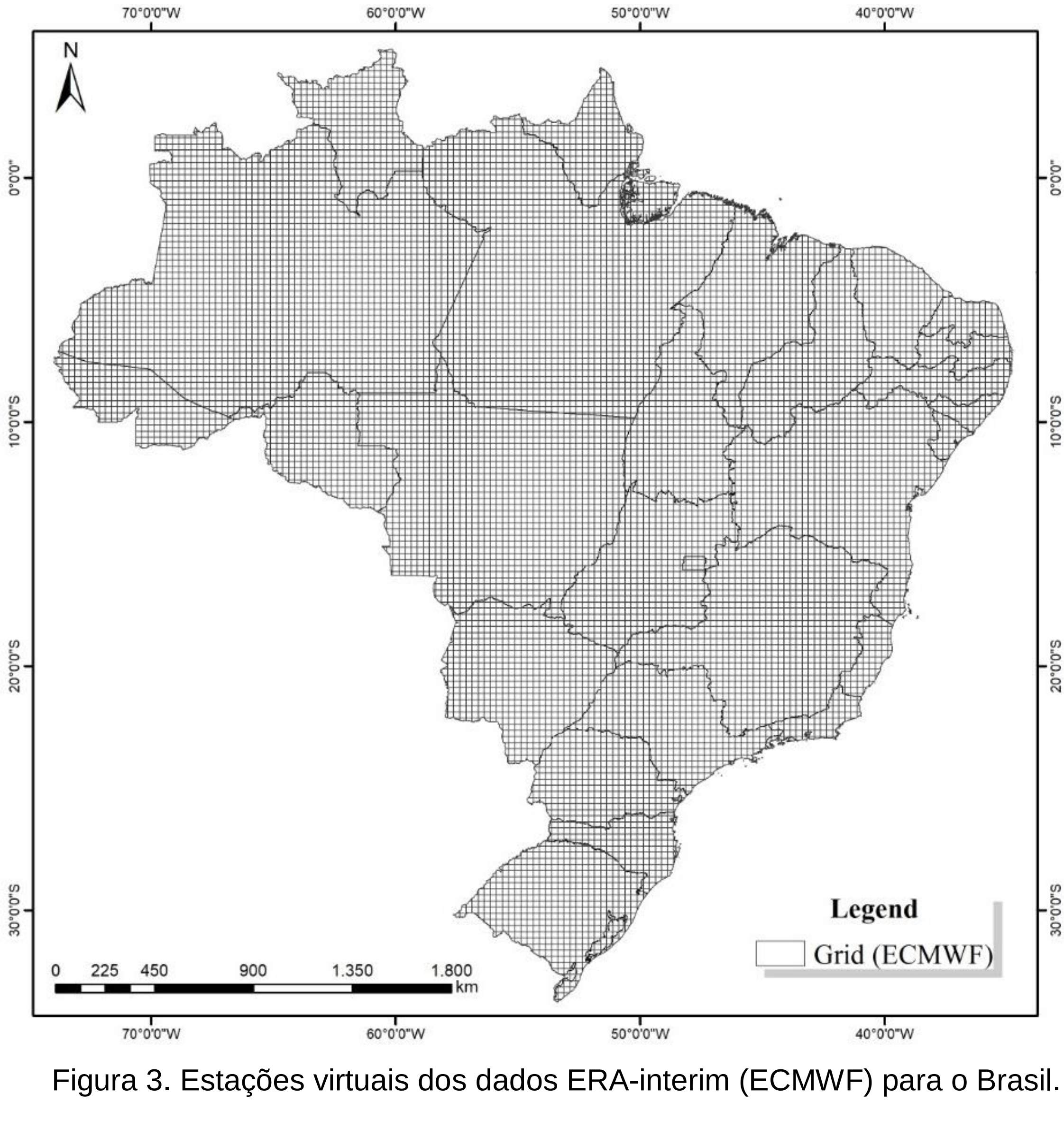

Figura 3. Estações virtuais dos dados ERA-interim (ECMWF) para o Brasil.

Guarnieri et al. (2007) utilizaram de GCM associado a redes neurais artificiais para realizar previsões de radiação solar. Para previsão de intensidade e volume de chuva, Vasconcellos et al. (2010), simularam razoavelmente bem as características da atmosfera assim como a precipitação máxima nas proximidades da Serra do Mar, na região Sudeste utilizando os GCM em alta resolução (espaçamentos horizontais de grade de 5 a 10 km). Não foram feitos ainda estudos de comparação de previsão de tempo e dados reais para áreas cafeeiras.

Os dados do ECMWF são largamente utilizados no mundo por apresentarem grandes vantagens em relação aos provenientes de estações meteorológicas de superfície, como rapidez na obtenção e ausência de valores faltantes (MORAES; ARRAES, 2012). O ECMWF tem o objetivo de auxiliar pesquisas científicas e melhorar a habilidade de realizar previsões, além de manter um arquivo de dados meteorológicos disponível gratuitamente na internet. Jung et al. (2012) demonstraram que as integrações de clima com os dados do ECMWF utilizando as resoluções horizontais, normalmente usado em previsões numéricas de tempo, levam a melhorias moderadas quando comparadas com a baixa resolução, pelo menos em aspectos de macroescala.

Na literatura são encontrados diversos trabalhos que demostram alta correlação entre os dados do ECMWF e os dados de superfície (DEPPE et al., 2006). Dentre esses trabalhos, podemos destacar Moraes e Arraes (2012) que utilizaram os dados do sistema ECMWF no estado do Paraná e observaram uma precisão de R^2=0,9 com os dados de superfície e o trabalho de Moraes et al. (2012) que calibraram os dados do ECMWF com dados de temperatura mínima e máxima do ar e precipitação de superfície no estado de São Paulo e observaram alta correlação.

A variabilidade dos elementos climáticos impacta consideravelmente na incidência das pragas e doenças da cultura do café. As pragas e doenças do cafeeiro atacam desde folhas, flores, frutos e até sementes, influenciando o desenvolvimento das plantas, proporcionando perdas na produção e na qualidade de bebida do café (MATIELLO et al., 2010). E poucos trabalhos tem buscado avaliar a influência das condições climáticas na incidência e severidades destas pragas e doenças.

1.4 Principais doenças e pragas do cafeeiro

As principais doenças do cafeeiro são a ferrugem do cafeeiro (*Hemileia vastatrix*) e a cercosporiose (*Cercospora coffeicola*), que podem causar elevados prejuízos na lavoura cafeeira. A ferrugem é uma doença fúngica que acarreta uma grande desfolha da planta (Figura 3). Nas regiões cafeeiras brasileiras onde as condições climáticas são favoráveis, elevados níveis de ferrugem causam perdas de 35% na produção (TEIXEIRA et al, 2007). A ferrugem é favorecida pelas condições climáticas de temperaturas do ar variando de 20 e 25ºC e precipitação total maior que 30 mm. Temperaturas do ar acima de 30ºC e abaixo de 15ºC são desfavoráveis à doença, no entanto, a epidemia da ferrugem aumenta rapidamente em temperaturas entre 15 e 18 ºC (PEREIRA et al., 2008). Diversos trabalhos demonstram que o grau de infecção da ferrugem do cafeeiro está diretamente relacionado com as condições climáticas (TALAMINI et al., 2003).

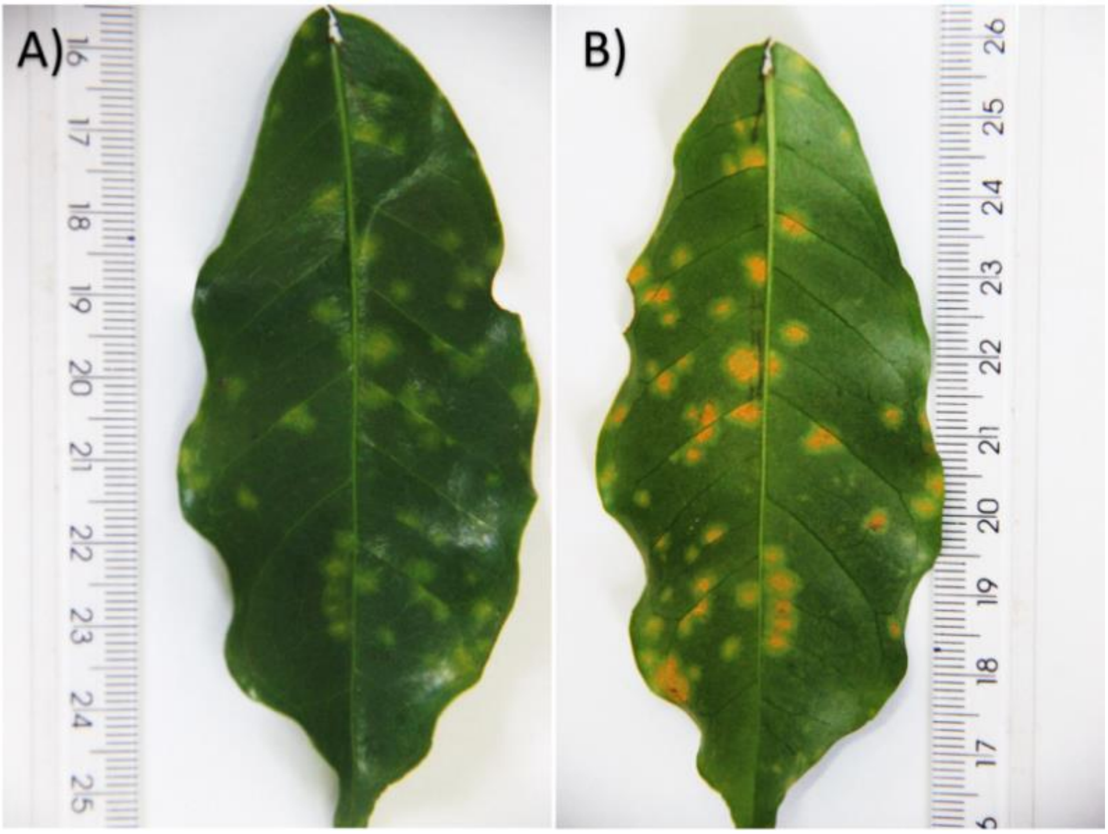

FIGURA 3. Infecção da doença de ferrugem do cafeeiro (*Hemileia vastatrix*). A) face adaxial (parte superior); B) face abaxial (parte inferior).

A cercosporiose infecta folhas (Figura 4) e frutos ocasionando desfolha e nos frutos provoca uma maturação precoce e queda prematura, aumentando o número de grãos chochos (CARVALHO et al., 2008). Os sintomas da cercosporiose no fruto ocorrem quatro meses após a floração, causando lesões deprimidas de cor castanho escuras, dispostas na direção do pedúnculo à coroa do fruto (POZZA et al., 2010; SANTOS et al., 2014). A utilização da irrigação no cultivo adensado proporciona um microclima favorável, favorecendo a cercosporiose do cafeeiro (PAIVA et al., 2013). Segundo Salgado et al., (2007) o déficit hídrico é uma das principais causas do aparecimento da cercosporiose.

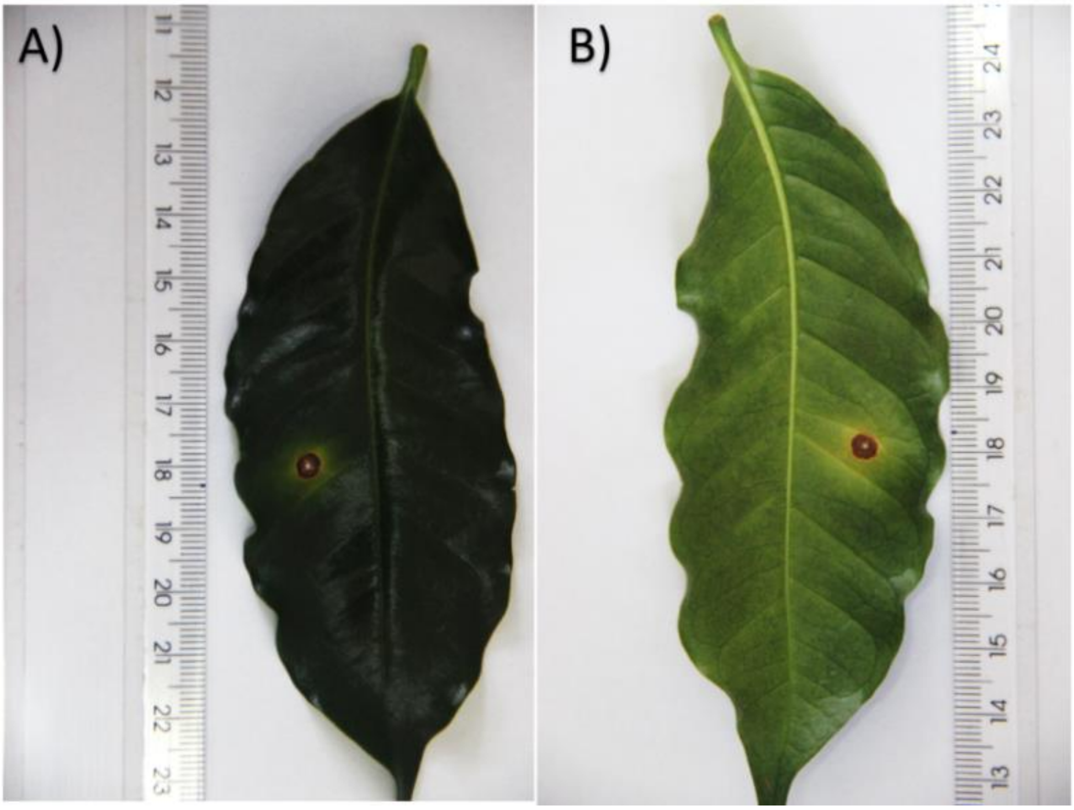

FIGURA 4. Infecção da cercosporiose (*Cercospora coffeicola*). A) face adaxial (parte superior); B) face abaxial (parte inferior).

As pragas contribuem para a queda da produtividade das lavouras e da qualidade do café produzido. A broca-do-café, Hypothenemus hampei (Coleoptera: Curculionidae, Scolytinae), e o bicho-mineiro, Leucoptera coffeella (Lepidoptera: Lyonetiidae), são as mais importantes pelos grandes prejuízos econômicos que causam ao reduzir a produtividade das lavouras e afetar a qualidade do café produzido (SOUZA et al., 2013).

O bicho-mineiro é considerado a principal praga-chave na atualidade (Figura 5), ocorrendo principalmente nas regiões de temperaturas do ar mais elevadas e com maior déficit hídrico (TEIXEIRA et al, 2007). O bicho-mineiro ocasiona grandes perdas ao cafeeiro, em virtude da derrubada das folhas o que promove a redução da capacidade fotossintética, chegando a proporcionar perdas de até 80% na produção (MENDONÇA et al., 2006).

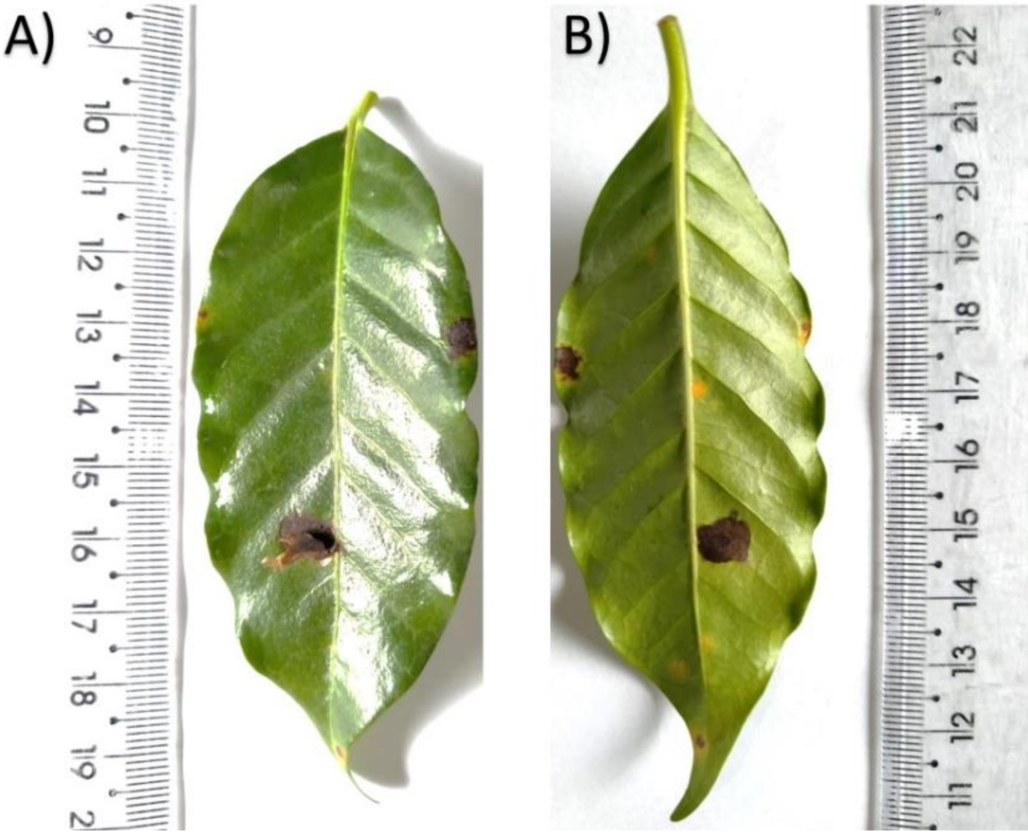

FIGURA 5. Infecção do bicho-mineiro (*Leucoptera coffeella*). A) face adaxial (parte superior); B) face abaxial (parte inferior).

A broca-do-café tem se dispersado por todas as regiões produtoras, atacando os países da Américas do Sul, Central e do Norte, África e Ásia (CANTOR et al., 2000). O ataque da broca-do-café causa prejuízo quantitativo (Figura 4), com a redução do peso dos grãos e queda de frutos, e prejuízo qualitativo, com a redução da qualidade do café através da alteração no tipo e bebida. Segundo Fernandes et al., (2014) os frutos são atacados em todos os estádios de crescimento comprometendo a produtividade e a qualidade da bebida.

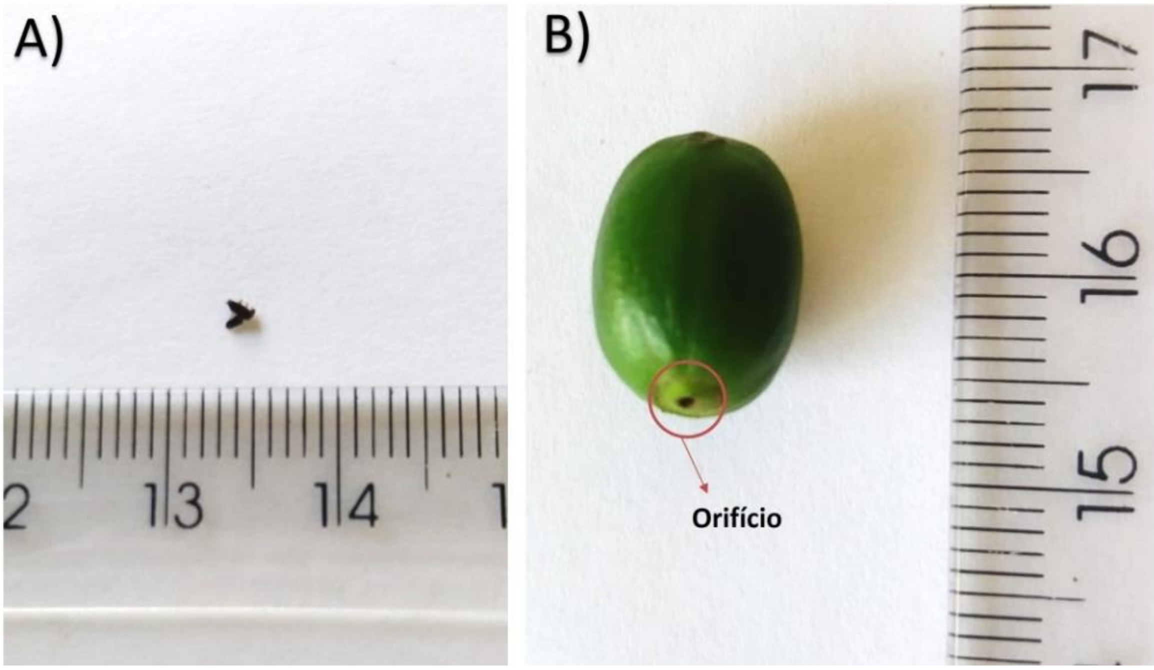

FIGURA 4. A) Inseto broca-do-café, *Hypothenemus hampei* (Coleoptera: Curculionidae, Scolytinae) e B) grão com orifício causado pela broca-do-café.

1.5 Algoritmos de machine learning

A estratégia mais comum de controle dessas doenças e pragas é a aplicação de agroquímicos, principalmente fungicidas e inseticidas foliares, dependendo da intensidade dos mesmos na região. Esse método tradicional deve ser melhorado uma vez que as questões ambientais está cada vez mais em pauta, principalmente com o aumento das certificações. A utilização de sistemas de alertas fitossanitários por meio de modelos de previsão dos índices de doenças e pragas.

Todas as relações existentes entre os elementos climáticas e a variabilidade das pragas e doenças do cafeeiro podem ser simulados com acurácia por meio de modelos agrometeorológicos (ROLIM et al., 2008) usando algoritmos de machine learning (SAHOO et al., 2017). Machine learning também conhecido como "aprendizagem de máquina" é um método que trabalha com análise de dados e busca automatizar a construção de modelos analíticos (SHEKOOFA et al.,, 2014; LI et al., 2016). É um campo da ciência da computação que trabalha com o reconhecimento de padrões utilizando da teoria do aprendizado computacional em inteligência artificial (SAHOO et al., 2017). Huber e Gillespie (1992) relatam que os algoritmos de machine learning podem automatizar os sistemas de alerta das doenças e pragas, pautando as tomadas de decisões dos produtores de quando é realmente necessário utilizar do controle químico nas lavouras.

As técnicas de machine learning são muitas promissoras para análises de *bigdata* mais rápidas, mais eficientes e acuradas (REHMAN et al., 2019). Os algoritmos machine learning utilizam de conceitos da teoria da probabilidade, estatística, teoria da decisão, otimização e técnicas de visualização, e por isso tem se tornado a vanguarda na área de *modelling* (SINGH et al., 2016; GÜMÜŞÇÜ et al., 2019).

Modelagem e o processo de desenvolvimento de um modelo. E modelo é a representação matemática simplificada de um sistema (JONES et al., 1987). Os modelos integralizam conhecimentos das áreas de agrometeorologia, fitossanidade, sensoriamento remoto, fisiologia vegetal, fitotecnia, ciência do solo e economia de forma interdisciplinar, podendo realizar estimações (JAME; CUTFORTH, 1996) e previsões (GOURANGA; ASHWANI, 2014) de variáveis. Com o uso de machine learning os computadores vão utilizar dos modelos e tomar as decisões com acurácia para os produtores.

Para Jame e Cutforth (1996) a modelagem auxilia na estratégia e tomadas de decisões nas cadeias produtivas, além de realizar simulações acuradas da dinâmica do crescimento de cultivos. A estimação é a quantificação de um fenômeno atual a partir de dados atuais (ROLIM et al., 2008). Nesta situação, ocorre a estimativa de uma variável sem ser antecipação, por exemplo, estimar a evapotranspiração do dia de hoje usando dados de temperatura do ar também de hoje.

Alguns exemplos de modelos de estimativa de doenças foram encontrados na literatura, como um modelo para Kushalappa et al. (1983) que propuseram explicar o curso de ação biológica de H. vastatrix, e com os resultados obtidos foram desenvolvimento de equações de regressão para avaliar a taxa de progresso da ferrugem-do-cafeeiro. Madden et al., (2000) que utilizaram de modelos para estimar as lesões da Plasmopara vitícola na videira em Ohio e assim melhorar o sistema de alerta e reduzir o número de aplicação de fungicidas. Yang et al., (2007) realizaram uma modelagem agrometeorológica para controle fitossanitário do pepino estimando a intensidade do míldio em ambiente protegido, ao final os autores observaram que as condições ideias para infecção foram uma amplitude térmica diária de 5°C, com temperatura média diária entre 15-25°C no outono e 80% de umidade relativa. E finalmente, Meira et al., (2009) que buscaram desenvolver árvores de decisão como modelos de alerta da ferrugem-do-cafeeiro em lavouras de café (*Coffea arabica* L.) com alta carga pendente de frutos, na localidade de Varginha, MG, observaram que a acurácia do modelo foi de 81% por validação cruzada, chegando a 89% segundo estimativa otimista, demonstrando ser uma alternativa viável às técnicas convencionais de aviso fitossanitário.

Na literatura são poucos os trabalhos de modelagem agrometeorológica que lidam com o processo de previsão, uma vez que a maioria dos trabalhos fazem estimações. A previsão é o processo de quantificação de um evento futuro a partir de dados disponíveis atuais (ROLIM et al., 2008). Neste caso, ocorre a previsão de uma variável com um determinado tempo de antecipação, por exemplo, 7 dias antes do fenômeno advir.

Alguns trabalhos que utilizam modelos para fins de previsão em sistemas agrometeorológicos de alerta fitossanitário, também conhecido como “disease early warning systems” foram encontrados na literatura. São exemplos desses: Kim et al., (2006) que utilizaram do processo de modelagem, com base na análise de árvore de regressão e lógica fuzzy, para prever a duração do molhamento (DPM) em cucurbitáceas, e ao final observaram uma simulação de 70% do DPM, melhorando assim o sistema de monitoramento. E, Baker et al., (2012) que buscaram desenvolver um modelo de previsão realizando previsão da requeima Batata na região de Great Lakes (EUA), conseguiram uma previsão acurada utilizando a metodologia de redes neurais artificiais. Entretanto, poucos trabalhos na literatura tem ousadia de prever a incidência de pragas e doenças do cafeeiro com dados climáticos usando algoritmos de machine learning.

Objetivo Geral

Este trabalho tem como objetivo prever a frequência de pragas e doenças do café e avaliar a influência do clima nessas enfermidades nas principais regiões produtoras de Minas Gerais.

Objetivos específicos

1. Avaliar quais os elementos meteorológicos, e em qual momento, exercem maior influência nas pragas: broca-do-cafeeiro e bicho-mineiro e doenças (ferrugem-do-cafeeiro e cercosporiose) do cafeeiro arábica nas principais localidades cafeeiras do Sul de Minas Gerais e do Cerrado Mineiro.
2. Desenvolver modelos agrometeorológicos para previsão de pragas e doenças em função das variáveis meteorológicas do Sul de Minas Gerais e Cerrado Mineiro usando algoritmos de machine learning com uma antecipação temporal suficiente para tomada de decisões.
3. Calibrar as variáveis meteorológicas: temperatura do ar e precipitação pluviométrica dos dados do ERA-Interim do centro ECMWF em relação aos dados de reais de superfície mensurados pelo sistema nacional de meteorologia (INMET) para o estado de Minas Gerais.

CAPÍTULO 2 – Models for simulating the frequency of pests and diseases of C*offea arabica* L.

ABSTRACT – Coffee is the most consumed beverage in the world, but phytosanitary problems are among the main causes of reduced productivity and quality. The application of foliar fungicides and insecticides is the most common strategy for controlling these diseases and pests, depending on their intensity in a region. This traditional method can be improved by using alert systems with models of disease and pest indices. We developed models for simulating trends over time as functions of the thermal index and models for estimating the levels of infestation of the coffee leaf miner and coffee berry borer and the severity of disease for coffee leaf rust and cercospora, the main phytosanitary problems in coffee crops around the world. We used historical series of climatic data and levels of pest infestation and disease severity in Coffea arabica for high and low yields for seven locations in the two main coffee-producing regions in the state of Minas Gerais in Brazil, Sul de Minas Gerais (Boa Esperança, Carmo de Minas, Muzambinho, and Varginha locations) and Cerrado Mineiro (Araxá, Araguari and Patrocínio locations), totalling 874 900 ha. We conducted two analyses. a) We simulated the trends of the progress of diseases and pests over time using non-linear models. Nonlinear logistic and Lorenz models were calibrated in function of the thermal index. We only used the thermal index because air temperature is commonly measured by farmers in the regions. b) We estimated the levels of pest infestation and disease severity using multiple linear regression, with the levels of diseases and pests as dependent variables and accumulated degree days (DD), coffee foliage (LF) estimated by DD and the number of nodes (NN) estimated by DD as independent variables. The best models (a) and estimates (b) were selected for accuracy using the mean absolute percentage error (MAPE), root mean square error, and an adjusted coefficient of determination (R^2adj). Cerrado Mineiro had the highest levels of pests and diseases, e.g. high-yielding coffee plantations in Araguari had intensities as high as 30.9% for rust, 36.1% for cercospora, 18.82% for the leaf miner and 4.5% for the berry borer, likely because Cerrado Mineiro averages 1 °C warmer than Sul de Minas Gerais. The difference between the dependent loads (high or low) of the coffee trees promoted different intensities only for rust, because the trees with high yields had higher intensities of the disease. The trend models for rust (FI) as a function of DD for Muzambinho (FIMUZ = f(DD)) for low yields and the models for the berry borer (BC) as a function of DD for Boa Esperança (BCBOE = f(DD)) for high yields were most accurate for Sul de Minas Gerais. The trend models for rust as a function of DD for Araxá (FIARX = f(DD)) for low yields and the models for cercospora (CI) as a function of DD for Patrocínio (CIPAT = f (DD)) for low yields were most accurate for Cerrado Mineiro. We used DD and LF = f(DD) and NN = f(DD) to predict diseases and pests with accuracy. MAPEs were 19.6, 5.7, 9.5, and 15.8% for rust, cercospora, leaf miner and berry borer, respectively, for Sul de Minas Gerais. Establishing phytosanitary alerts using only air temperature was possible with these models.

KEY-WORDS: crop model; early prevision; phytosanitary alerts; forecasting

Introduction

Brazil is the largest producer and exporter of coffee in the world (Meinhart *et al.,* 2017), with a total of 2 million ha planted (CONAB, 2019). Production areas in Brazil are distributed in the south-central region, in the states of Minas Gerais, São Paulo, and Espírito Santo (Andrade *et al.,* 2012; Cubry *et al.,* 2013). Minas Gerais represents about 50% of the total production of Brazil and is the largest producer of coffee in the country (Barbosa *et al.,* 2012, Ronchi *et al.,* 2015).

Coffee leaf rust (*Hemileia vastatrix*) and cercospora (*Cercospora coffeicola*) are the most common phytosanitary problems in coffee-producing regions on all continents (Haddad *et al.,* 2009; Ghini *et al.,* 2015, Castro *et al.,* 2018), which can cause high losses in coffee plantations. Coffee leaf rust is one of the main diseases affecting coffee plantations worldwide, having an important economic impact in the coffee industry in countries where coffee is an important part of the economy (Esquivel, Sanchez, Barbosa, 2017). Rust is a fungal disease that causes severe defoliation of plants. High levels of rust cause production losses of 35% in the Brazilian coffee regions where climatic conditions are favourable (Teixeira *et al.,* 2007).

Rust has also been reported in Peru (Castro *et al.,* 2018), Colombia, and Central America (Esquivel, Sanchez, Barbosa, 2017), and cercospora has also been reported in Thailand (Sirinunta and Akarapisan, 2015) and Latin America (Ghini *et al.,* 2015), where both cause heavy losses. These diseases cause worldwide economic losses of approximately 1 billion $US, due to high mitigation costs, and decreases in productivity of up to 50% (Haddad *et al.,* 2009; Jackson *et al.,* 2012).

Cercospora infects leaves and fruits, causing defoliation, early fruit maturation, and premature fruit fall, increasing the number of pimples (Carvalho *et al.,* 2008). The symptoms of cercosporiosis in the fruit occur four months after flowering, causing dark-brown depressions from the peduncle to the crown of the fruit (Santos *et al.,* 2010). Irrigating dense crops provides a microclimate favouring cercosporiosis in coffee (Paiva *et al.,* 2013). Water deficits are one of the main causes of the appearance of cercosporiosis (Salgado *et al.,* 2007).

Pests also contribute to losses of productivity and the quality of coffee. The main pests are the coffee berry borer (*Hypothenemus hampei*) (Coleoptera: Curculionidae, Scolytinae) and the coffee leaf miner (*Leucoptera coffeella*) (Lepidoptera: Lyonetiidae) due to the large economic losses they cause (Souza *et al.,* 2013). The berry borer is dispersed throughout the coffee-producing regions of the world, attacking countries in Africa, Asia, and South, Central, and North America (Cantor *et al.,* 2000). The attack of the berry borer causes quantitative damage, reducing grain weight and increasing fruit drop. The fruits are attacked at all stages of growth, compromising the productivity and quality of the beverage (Fernandes *et al.,* 2014). The leaf miner is currently considered the main pest in regions with high air temperatures and water deficits (Teixeira *et al.,* 2007). The leaf miner causes large losses to coffee trees, by reducing the photosynthetic area, and leaching losses of up to 80% of production (Mendonça *et al.,* 2006). Leaf miners have been reported in several other countries, such as Costa Rica (Allinne *et al.,* 2016) and Ethiopia (Abedeta *et al.,* 2015).

Models are simplified mathematical representations of a system, and their development is called modelling (Jones *et al.,* 1987). Models integrate information from areas such as agrometeorology, phytosanitation, remote sensing, plant physiology, plant science, soil science, and economics for estimating (Jame and Cutforth, 1996) and predicting (Gouranga and Ashwani, 2014) a variety of parameters, assisting in the formation of strategies and decisions in production chains, and accurately simulating the dynamics of crop growth (Jame and Cutforth, 1996).

We used non-linear models because the adjusted parameters we used were physically and/or biologically relevant to the system under study (Gujarati and Porter, 2001). Nonlinear models have been used in other studies. For example, Kin *et al.* (2001) estimated the developmental times of *Carposina sasakii* (Lepidoptera: Carposinidae) as a function of the average air temperature in the Suwon region of Korea. The authors observed that the models were accurate and that the times decreased with increasing temperature up to 32 °C in the eggs, 28 °C in the larvae, and 30 °C in the pupae. And, Rowley et al. (2017) who predicted the development of *Haplodiplosis marginata* (Diptera: Cecidomyiidae) in the United Kingdom using non-linear models (a binomial generalised linear model and Weibull and Probit models), the pest had a sigmoidal tendency of growth up to 1550 accumulated degree days (DD), and the growth curve stabilised at an accumulation of 980 DD.

Diseases and pests decrease coffee quality and productivity (Spongoski *et al.,* 2005). Rates of disease and pest infestation in coffee plants have been reported for Brazil and all countries where the plant is cultivated. The application of foliar fungicides and insecticides is the most common strategy to control these diseases and pests, depending on their regional intensity. This traditional method can be improved using alert systems with models of disease and pest indices. We thus sought to develop models for simulating of the growth curve of pests and diseases over time as a function of the thermal index and estimating pest infestation and disease severity of *Coffea arabica* L. for use in phytosanitary risk-alert systems.

Material and methods

We used a historical series of climatic data and levels of pest infestation and disease severity of arabica coffee for the state of Minas Gerais, Brazil. Representative sites of coffee production were Boa Esperança (BOE), Carmo de Minas (CDM), Muzambinho (MUZ), and Varginha (VAR) in Sul de Minas Gerais (SO_{MG}) and Araxá (ARX), Araguari (ARG) and Patrocínio (PAT) in Cerrado Mineiro (CE_{MG}) (Table 1 and Fig. 1).

TABLE 1. Geographic characteristics of the main coffee-growing areas of the state of Minas Gerais.

Locations	Latitude (º)	Longitude (º)	Altitude (m)	Period	Area (km_2)	
				Sul de Minas Gerais		
Boa Esperança	21º 03' 59" S	45º 34' 37" W	830	2010 to 2018	860,7	-
Carmo de Minas	22º 10' 31" S	45º 09' 03" W	1080	2006 to 2018	323,3	-
Muzambinho	21º 20' 47" S	46º 32' 04" W	1033	2010 to 2018	409	-
Varginha	21º 34' 00" S	45º 24' 22" W	940	1998 to 2018	395,6	-
				Cerrado Mineiro (Alto Paranaíba)		
Araguari	18º 59' 35" S	46º 59' 01" W	961	2010 to 2018	2.731	-
Araxá	19º 33' 21" S	46º 58' 08" W	960	2010 to 2018	1.165	-
Patrocínio	18º 33' 21" S	48º 12' 25" W	933	2010 to 2018	2 866	-

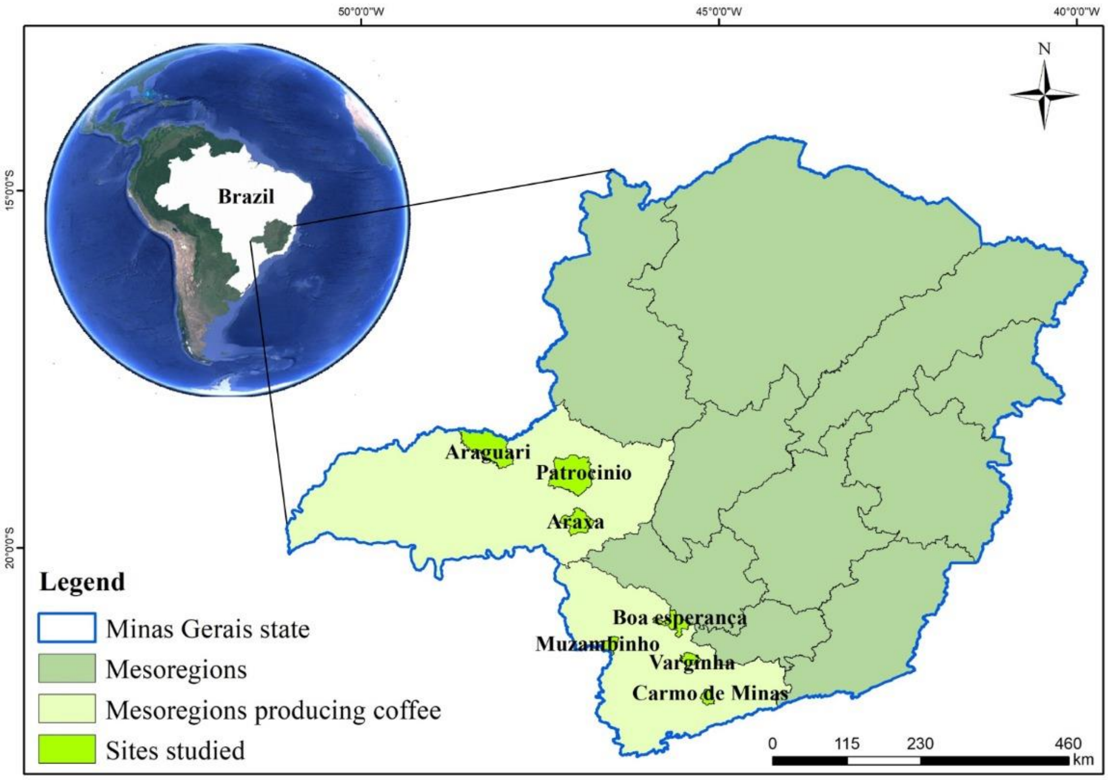

FIGURE 1. Coffee regions in the state of Minas Gerais, Brazil, analysed in this study.

Daily meteorological data were obtained from automatic meteorological stations installed near the coffee plantations evaluated: air temperature (T_{AIR}) and precipitation (P). Data were collected using a Meteorological Station Vantage Pro2 Davis-K6162 (*Davis* Instruments, Hayward, Californian-USA) data acquisition system. The T_{AIR} and P data were used to estimate reference potential evapotranspiration (PET) following the method of Camargo (1971) (Eq. 1). The availability of data was the criterion for choosing this model.

$$PET = 0.01 \cdot \left(\frac{Qo}{2.45}\right) \cdot Tair \cdot ND \qquad (1)$$

where Qo is the solar irradiance at the top of the atmosphere (MJ m^{-2} d^{-1}), T_{AIR} is the mean air temperature (°C), and ND is the number of days in the period.

We estimated the components of the water balance (WB) proposed by Thornthwaite and Mather (1955) using an available water capacity (CAD) of 100 mm, because this value represents the majority of soils in the main coffee regions (Meireles *et al.,* 2009). The WB components were: soil-water storage, water deficit (DEF), and surplus water (EXC) of the soil-plant-atmosphere system.

The data for disease severity and pest infestation were provided by the Procafé Foundation (http://fundacaoprocafe.com.br/) from field evaluations with no phytosanitary treatment at the sites (Table 1). The data were for the diseases coffee rust (*H. vastatrix*) and cercospora (*C. coffeicola*) and the pests coffee leaf miner (*L. coffeella*) and coffee berry borer (*H. hampei*). These data had been collected from fields with steep and gentle slopes, also for 1995-2018.

Incidences were measured non-destructively. The plants were randomly chosen in a zigzag walking pattern in the area, as recommended by (Chalfoun 1997). Incidence scores were determined for leaves from the third or fourth knot of branches at the middle third of the plants. Detailed methodology is presented in Table 2.

TABLE 2. Methodology for evaluating coffee development, diseases, and pests used by the Procafé Foundation of Brazil.

Phytosanitary problem	Methodology
Phenology	
Number of nodes	- Sample 20 plants per plot (random) - Select four branches per plant in the middle third (one on each side) - Quantify the number of nodes developed from September in each chosen branch
Leafiness of coffee plant	- Sample 20 plants per plot (random) - Select four branches per plant in the middle third (one on each side) - Quantify the amount of foliar development in the chosen branches - Quantify the soil-foliage percentages in the samples
Diseases	
Coffee rust (*Hemileia vastatrix*) and Cercospora index (*Cercospora coffeicola*)	- Sample 20 plants per field - Collect the leaves in the middle third of the chosen plant - Choose five lateral branches at random on each side of the plant - Remove a completely developed leaf, of the 3rd or 4th pair of leaves, from each branch - Total of 10 leaves/plant (five on each side) - Total of 200 leaves/field - Quantify the percentage of disease in the samples
Pests	
Coffee leaf miner (*Leucoptera coffeella*)	- Sample 20 plants per field - Collect the leaves in the middle third of the chosen plant - Choose five lateral branches at random on each side of the plant - Remove a completely developed leaf, of the 3rd or 4th pair of leaves, from each branch - Total of 10 leaves/plant (five on each side) - Total of 200 leaves/field - Quantify the percentage of the pests in the samples
Coffee berry borer (*Hypothenemus hampei)*	- Sample average of 50 plants per plot - Choose four branches per plant (one on each side) - Collect 25 fruits/branch for a total of 100 fruits/plant - 50 plants/field for a total of 5000 fruits/field - Quantify the percentage of the pest in the samples

Each node normally develops from the plagiotropic branch of the coffee tree to have two leaves (one on each side of the node). Leafiness and the percentages of diseases and pests were quantified by:

$$Leafiness\ (\%)\ =\ [\frac{(\frac{Leaf\ number}{2})}{Number\ of\ nodes}]\ \cdot\ 100 \qquad (2)$$

$$CLM\ (\%)\ =\ [\frac{(Number\ of\ leaves\ with\ live\ larvae)}{Number\ of\ leafs\ total}]\ \cdot\ 100 \qquad (3)$$

$$CBB\ (\%)\ =\ [\frac{(Number\ of\ fruits\ with\ berry\ borer)}{Number\ of\ fruits\ total}]\ \cdot\ 100 \qquad (4)$$

$$FI\ (\%)\ =\ [\frac{(Number\ of\ leaves\ with\ Ferrugem)}{Number\ of\ leafs\ total}]\ \cdot\ 100 \qquad (5)$$

$$CI\ (\%)\ =\ [\frac{(Number\ of\ leaves\ with\ Cercospora)}{Number\ of\ leafs\ total}]\ \cdot\ 100 \qquad (6)$$

where Ci is the cercospora index (%), FI is coffee rust (%), CBB is the coffee berry borer (%), and CLM is the leaf miner (%).

The diseases and pests of the coffee plantations were evaluated for high and low productivity, which occur due to the natural bienniality of coffee plants. High productivity represents >30 bags (60 kg each) ha^{-1}, and low productivity represents <10 bags ha^{-1}. The discontinuous range of this classification is due to the field differences for subsequently high and low yields. The cultivars were Catuaí and Mundo Novo, both susceptible to the diseases and pests.

Pest infestation, disease severity, and the time of each assessment were used to construct trend curves for describing the development of pest infestation and disease severity over time (Nutter, 1997; Bergamin Filho, 2011). We analysed the trend of development of the diseases and pests of coffee trees as the DDs for high and low yields for all sites. Total DD was calculated using a base temperature (Tbase) of 10.2 °C, as proposed by Carvalho et al. (2014) (Eq. 7) for the development of coffee crops.

$$\sum DD = \left[\frac{Tmax+Tmin}{2}\right] - Tbase \qquad (7)$$

where ΣDD is total DD after the beginning of vegetative growth standardised from September, Tmax is the absolute maximum daily air temperature (°C), and Tmin is the absolute minimum daily air temperature (°C).

The development of disease over time was simulated using non-linear models. Growth models were adjusted by a regression analysis using sigmoid models with four parameters (logistic) and the Lorentz model with four parameters (Eqs. 8 and 9, respectively). The normal distribution of the deviations was verified using the Kolmogorov-Smirnov adhesion test at P<0.05.

$$y = y_{max} + \frac{Ymax - Ymin}{1+(\frac{x}{x_0})^p} \qquad (8)$$

$$y = y_{min} + \frac{2 \cdot A}{\pi} \frac{w}{4\,(x-x_0)^2+w^2} \qquad (9)$$

where Ymax is maximum development, Ymin is minimum development, X_0 is the number of degree days for maximum disease development, p is the maximum growth rate at X_0, W is the midpoint standard deviation, and x is DD after the beginning of vegetative growth, always from 01 September, as suggested by Hinnah et al. (2018).

The parameters of all nonlinear models were estimated by the least-squares method, which consists of minimising the sum of the squares of the model errors and iteratively using linear programming with the generalised reduced gradient model, as proposed by Lasdon and Waren (1982). We generated the confidence interval (IC) of the estimate (at 5% significance) for all models for estimating trends, as proposed by Neyman (1937) (Eq. 10). Confidence intervals are indicators of precision and the stability of an estimate, which is an evaluation of how close a measurement is to the original estimate:

$$IC = Y_{est} \pm TableT_{(\alpha/2, n-2)} \times S_{Yest} \times \sqrt{\frac{1}{n} + \frac{(x_{obs} - \bar{x})^2}{(n-1) \times S_X^2}} \qquad (10)$$

where Y_{est} is a value estimated by the model, s_{Yest} is the standard deviation of the values estimated by the model, X_{obs} is an observed value of disease severity or pest infestation, $\bar{x}$ is the mean of these observed values, s^2 is the variance of the data, n is the number of datapoints, and TableT (α/2,n-2) is value found from the table of the normal distribution (α = 0.05).

The models for estimating levels of disease severity and pest infestation were based on multiple linear regressions. The dependent variable was the levels pests and diseases and the independent variables were the NN and tree foliage accumulated in DD (Eq. 11). One model for each disease and pest was calibrated for the entire region using BOE, CDM, MUZ, and VAR data for SO_{MG} and ARG, ARX, and PAT data for CE_{MG}. Only models with *P*<0.05 were selected:

$$Y = \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \beta_3 \cdot X_3 + \cdots + \beta_o \qquad (11)$$

where *Y* is the level of pest infestation or disease severity, β_N is the adjusted weight, X_N is an independent variable, and β_O is an linear coefficient.

The actual observed field data and the results of all models were compared using several statistical indices: accuracy, precision, and level of significance (Table 3). Accuracy indicates the closeness of an estimate to the observed value and was evaluated using the mean absolute percentage error (MAPE) and the root mean square error (RMSE). Precision is the ability of a model to repeat an estimate and was evaluated using the coefficient of determination (R^2) adjusted (R^2adj) following Cornell and Berger (1987).

TABLE 3. Accuracy and precision of statistical indices used for evaluating the non-linear adjustments of rates of disease severity and pest infestation. Yest, estimated value of y; Yobs, observed value of y; Xobs, observed value of x; n, number of datapoints. Overlined variables are means.

Statistical Index	**Equation**
Accuracy	
MAPE	$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left(\left\lvert\frac{Yest_i - Yobs_i}{Yobs_i}\right\rvert \times 100\right)$
RMSE	$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(Y_{OBS_i} - Y_{EST_i}\right)^2}{N}}$
Precision	
R^2	$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(Yest_i - \overline{Y}obs\right)^2}{\sum_{i=1}^{n}\left(Yest_i - \overline{Y}obs\right)^2 - \sum_{i=1}^{n}\left(Yest_i - Yobs_i\right)^2}$

Results and discussion

Coffee production in SO_{MG} and CE_{MG} is characterised by the presence of diseases and pests, which is strongly influenced by climatic conditions. SO_{MG} and CE_{MG} had similar climatic trends (Figure 2), e.g. higher T_{AIR} and P of 300DD-2800DD and lower T_{AIR} and smaller P after the accumulation of 2800 DD. DEF was highest in two periods, from 0 DD to 1000 DD and from 2900 DD to 4000 DD in both regions (Figure 2AB).

The accumulation of P, PET, EXC, and DEF was higher for CE_{MG}, at 1446.84, 871.64, 771.84, and -180.61 mm, respectively. The accumulation of P, PET, EXC, and DEF in SO_{MG} was 1380.21, 859.49, 614.15, and -105.60 mm, respectively (Figure 2C, D). The lower DEF in SO_{MG} has previously been reported (e.g. Barbosa et al., 2010), who found that DEF was usually higher in CE_{MG}.

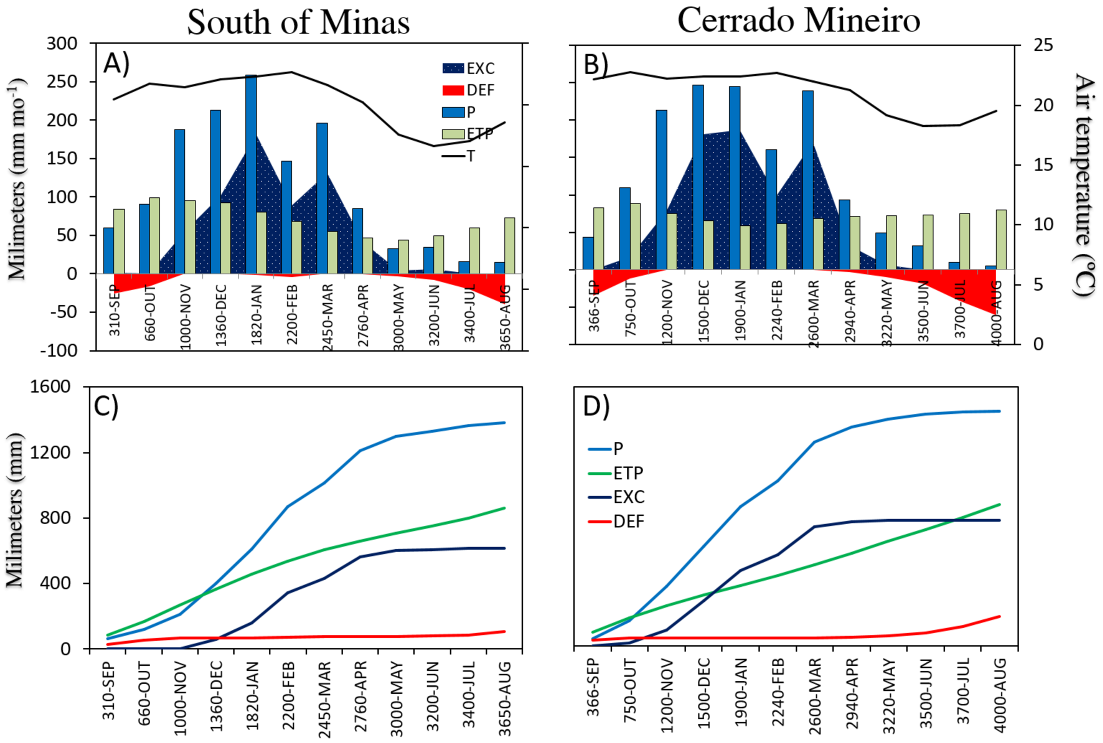

FIGURE 2. Seasonal variation of climatic parameters (A and B) and accumulation of climatic parameters after the resumption of vegetative growth (C and D) in Sul de Minas Gerais and Cerrado Mineiro.

The average rust severity for high and low yields was 33.5 and 15.8% in SO_{MG} (Figure 3A) and 27.1 and 24.9% in CE_{MG} (Figure 3B), respectively. Coffee trees in SO_{MG} with low yields tended to have less rust than trees with high yields, because coffee plants with high yields allocate photoassimilates for production (Aparecido *et al.,* 2018), so the plants are more vulnerable to rust.

Disease rates were higher in CE_{MG} than SO_{MG} (Figure 3). Carvalho *et al.* (2017) also reported a higher incidence of rust and cercospora in CE_{MG} than SO_{MG}. Levels were highest for ARG; 30.9% of high-yield trees had rust, 36.1% had cercospora, 18.82% had leaf miners, and 4.5% had berry borers (Figure 3B, D, F, H). In contrast, levels were lowest for VAR; 20.9% of low-yield trees had rust, 2.7% had cercospora, 1.38% had leaf miners, and 0.18% had berry borers (Figure 3A, C, E, G).

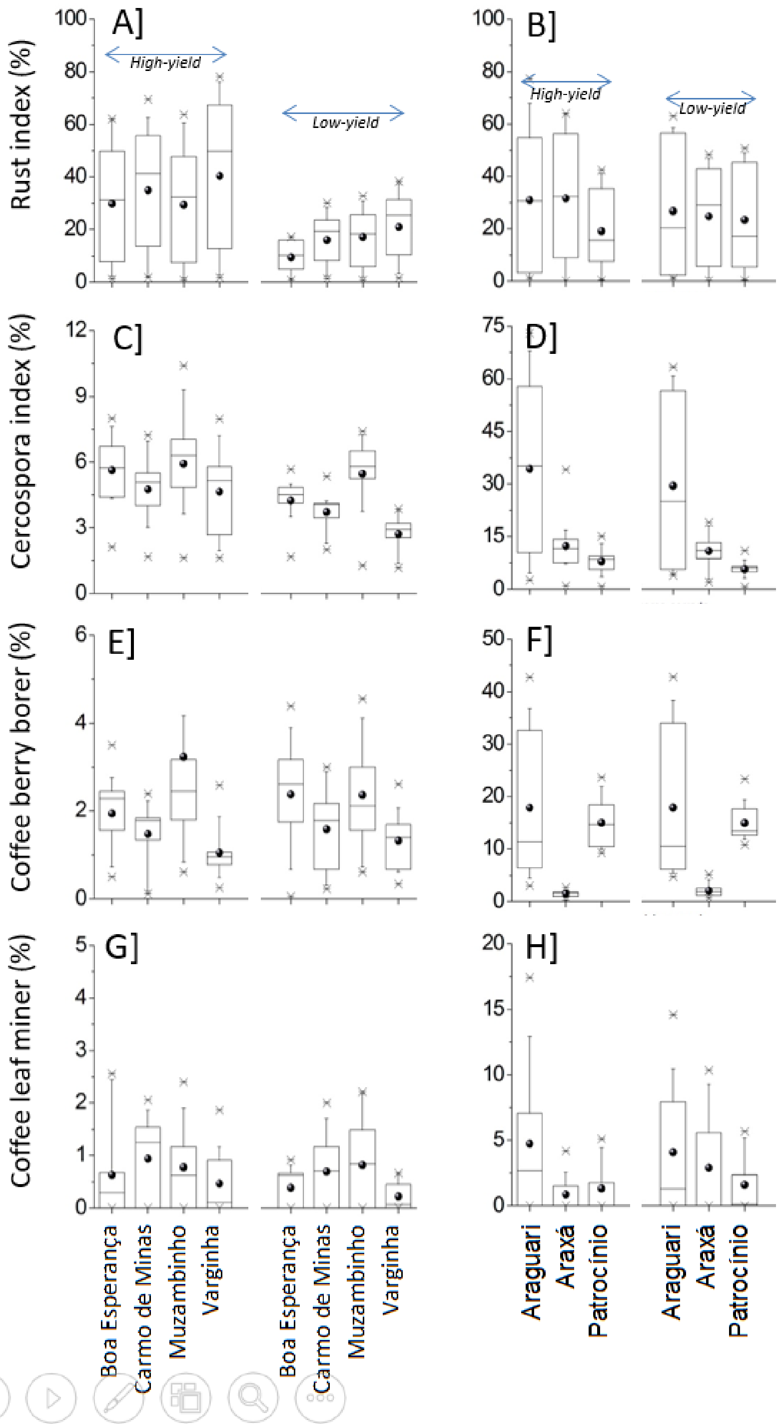

FIGURE 3. Box plot: A and B) Coffee leaf rust index, C and D) cercospora index, E and F) coffee berry borer, and G and H) coffee leaf miner for high and low yields for Sul de Minas Gerais (Boa Esperança, Carmo de Minas, Muzambinho and Varginha), and Cerrado Mineiro (Araguari, Araxá and Patrocínio) for 1995-2015. Legend: •, average; ─, median; □, 50% of values; I, 99% of values; X, extreme values (outliers).

The analysis of the growth trends based on DD identified several relationships. The growth of NN per plagiotropic branch due to the accumulation of DD demonstrated a hyperbolic adjustment (Figure 4). The maximum value of NN was eight nodes per branch, at 3850 DD in SO_{MG} and 10 nodes per branch at 4000 DD in CE_{MG}, for BOA and ARG, respectively. NN per plagiotropic branch was positively correlated with coffee productivity (Assis *et al.,* 2014), because these nodes produced coffee fruits.

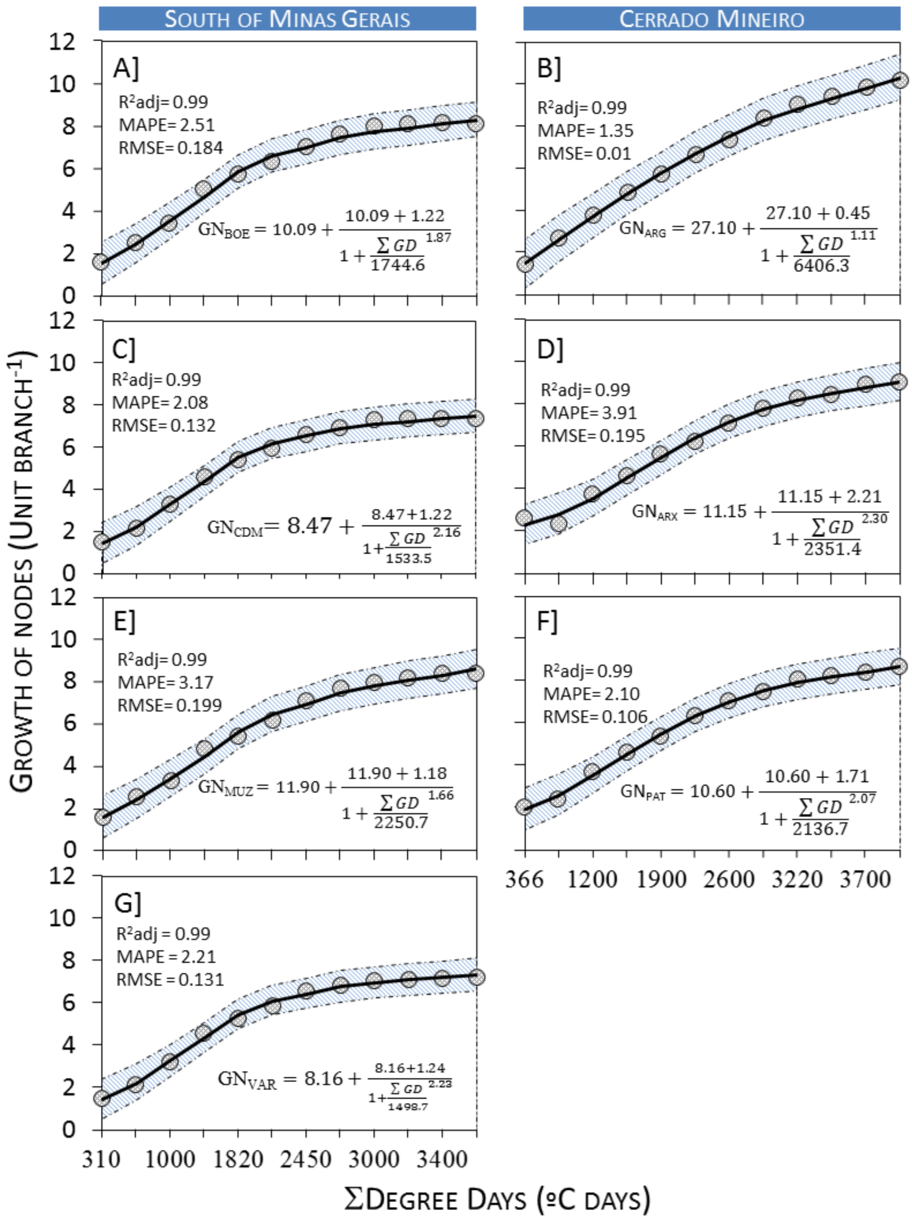

FIGURE 4. Growth of nodes (NN) based on the total degree days (ΣDD, °C d^{-1}) after the beginning of vegetative growth (September). A) Boa Esperança, B) Araguari, C) Carmo de Minas, D) Araxá, E) Muzambinho, F) Patrocínio, and G) Varginha.

Leafiness based on DD was similar in SO_{MG} and CE_{MG}. The reduction of foliage in SO_{MG} occurred from 2400 DD but was after 2500 DD in CE_{MG} (Figure 5). This early defoliation was attributed to the pests and foliar diseases, which reduce the photosynthetic area of plants and may even promote the death of plagiotropic branches (Custodio *et al.,* 2010; López-Duque, Fernández-Borrero, 1969; Gree, 1993).

Leafiness was lowest in SO_{MG} at 18.37% and in CE_{MG} at 17.1%, in CDM and ARX, respectively (Figure 5C, D). Foliar density is directly linked to crop production, because leaves are one of the main organs that contain chloroplasts, the organelle entirely responsible for photosynthesis (Taiz and Zeiger, 2009).

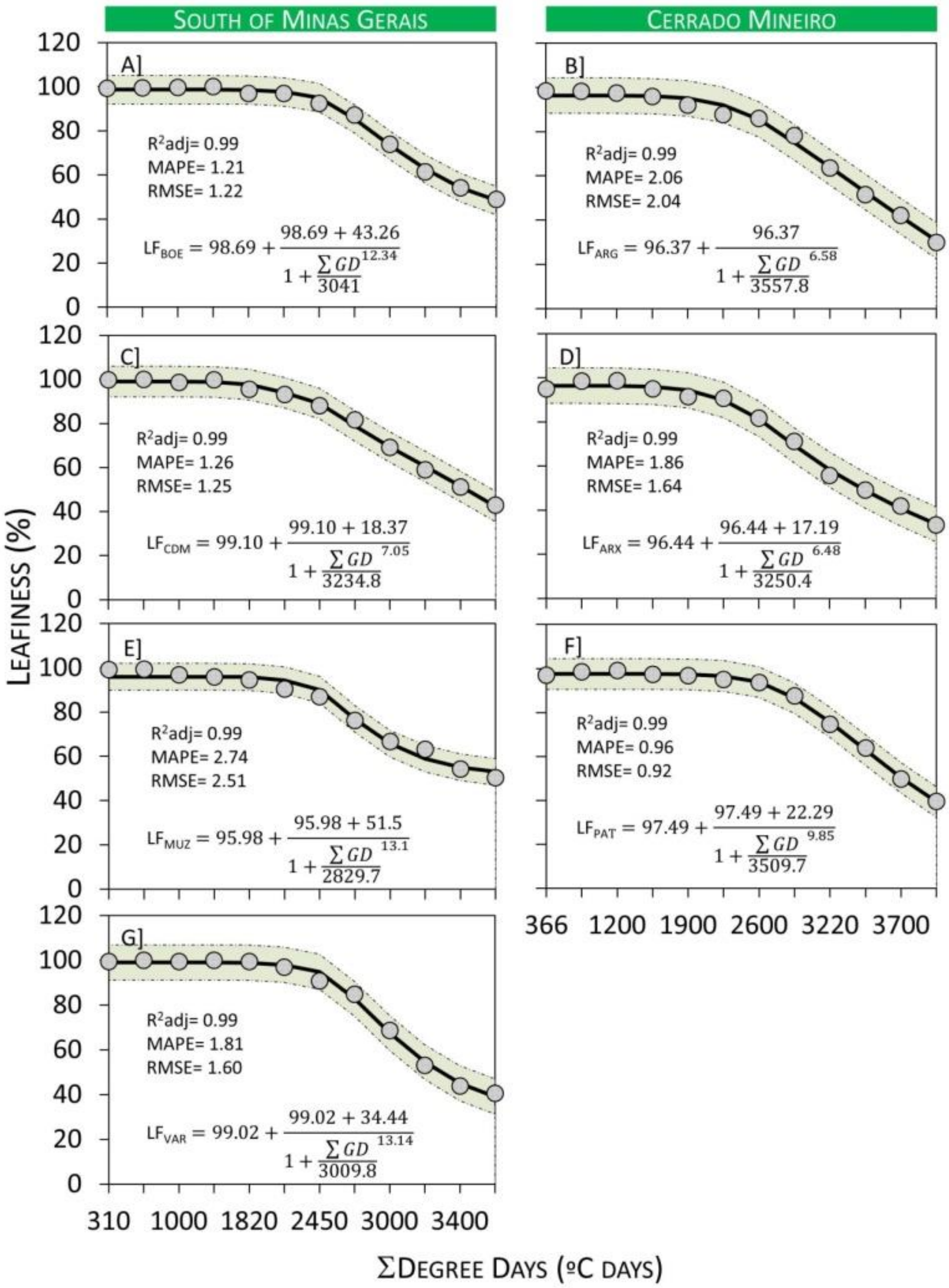

FIGURE 5. Leafiness based on total degree days (ΣDD, °C d^{-1}). A) Boa Esperança, B) Araguari, C) Carmo de Minas, D) Araxá, E) Muzambinho, F) Patrocínio, and G) Varginha. Points are observed values, solid lines are estimated values, and dashed lines indicate 95% confidence intervals.

Coffee rust based on DD had a sigmoidal adjustment (Figures 6 and 7). The rust index in SO_{MG} was higher for high than low yields (Figure 6). The rust index was more similar in CE_{MG} for both yields (Figure 7). Meira *et al.* (2009) reported that trees were more predisposed to rust in years of high than low yields, because the trees allocate more photoassimilates for the production of high yields (Aparecido *et al.,* 2018). Carvalho *et al.* (2001) evaluated the rust indices as a function of the level of coffee production and found that increasing levels of production were correlated with an increase in the susceptibility of the plants to rust.

The adjusted models were all significant ($P<0.001$), with low MAPEs and RMSEs and with R^2adj ≈1.0, indicating that the rust index could be estimated as a function of DD. For example, the logistic model estimating the rust index for BOE (FI_{BOE}) for high yields had a MAPE of 9.91%, RMSE of 2.80, and R^2adj of 0.98 (Figure 6A). An error of 9.91% is low, because a rate of maximum disease severity of 72.27% at the end of the cycle has a variation of only ±7.14%, which is low for a disease. The models with the highest accuracy in SO_{MG} for estimating rust for the high yields were $\Sigma FI_{BOE} = F(\Sigma DD)$ and $\Sigma FI_{CDM} = F(\Sigma DD)$, with MAPEs of 9.91 and 15.01%, respectively. For the crop adjustments for low yields, $\Sigma FI_{MUZ} = F(\Sigma DD)$ and $\Sigma FI_{BOE} = F(\Sigma DD)$ were the best models, with MAPEs of 8.94 and 12.05%, respectively.

The parameter p adjusted in the sigmoid models indicates the maximum rates of growth of the diseases. The rust index for BOE, CDM, MUZ, and VAR had maximum rates of 5.60, 4.36, 3.25, and 6.08% for high yields and 6.75, 3.54, 3.82, and 3.41% for low yields, respectively (Figure 6). The highest rate of rust growth in SO_{MG} was 6.75% °C^{-1} d^{-1}. This rate was for a tree with a low yield in BOE at 2484.5 °C d^{-1} (Figure 6B). The lowest of the highest growth rates was 3.25% °C^{-1} d^{-1} in MUZ at 2792 °C d^{-1} in a high-yield coffee tree (Figure 6E). The highest rates of coffee development in SO_{MG} for high yields occurred at 2541.7, 2418, 2792, and 2215.3 °C d^{-1} for BOE, CDM, MUZ, and VAR, respectively. The highest growth rates for low yields were earlier, at 2484.5, 2263.6, 2210.8, and 2246 °C d^{-1} for BOE, CDM, MUZ, and VAR, respectively (Figure 6).

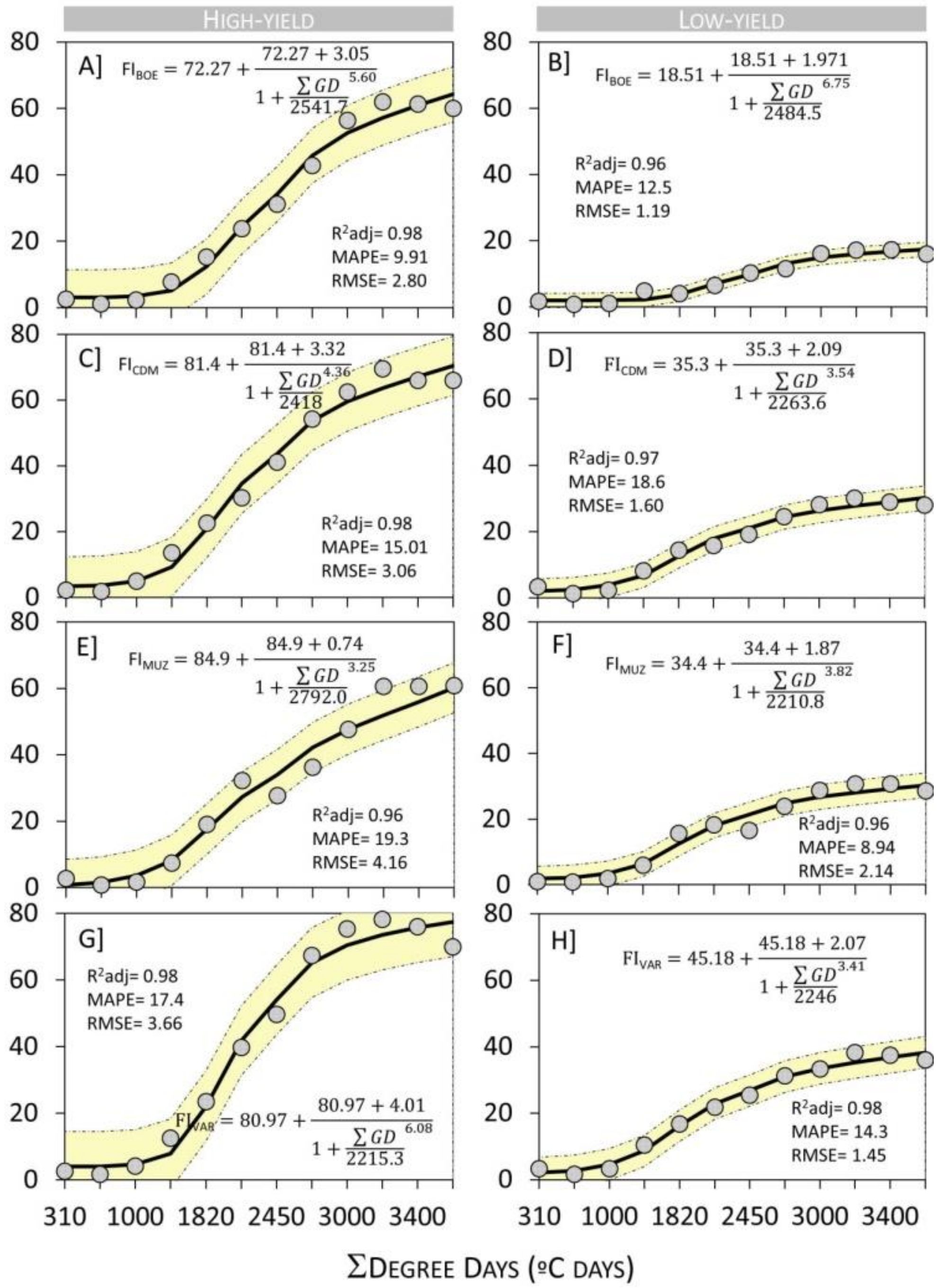

FIGURE 6. Coffee rust index (FI) based on total degree days (ΣDD, °C d^{-1}) after the beginning of vegetative growth (September) in Sul de Minas Gerais. A and B, Boa Esperança; C and D, Carmo de Minas; E and F, Muzambinho; and G and H, Varginha. Points are observed values, solid lines are estimated values, and dashed lines indicate 95% confidence intervals.

The models for CE_{MG} with the highest accuracy for estimating rust for high yields were $\Sigma FI_{ARG} = F(\Sigma DD)$ and $\Sigma FI_{ARX} = F(\Sigma DD)$, with MAPEs of 12.4 and 15.2%, respectively (Figure 7A, C). The best models for the crop adjustments for low yields were, in ascending order, $\Sigma FI_{ARX} = F(\Sigma DD)$ and $\Sigma FI_{ARG} = F(\Sigma DD)$, with MAPEs of 11.8 and 14.4%, respectively (Figure 7B, D). The rust index for CE_{MG} plants had maximum rates of 10.1, 6.09, and 16.4% (high yields) and 9.87, 6.61, and 8.46% (low yields) in ARG, ARX, and PAT, respectively (Figure 7). These rates were higher than those for rust in SO_{MG} (Figure 6).

The highest growth rate was 16.4% 1 °C^{-1} d^{-1} for rust, much higher than for SO_{MG} (6.75% °C^{-1} d^{-1}). The growth rate of 16.4% °C^{-1} d^{-1} occurred in a tree with a high yield in PAT, at 2705.6 °C d^{-1} (Figure 7E). The highest rates of coffee development in CE_{MG} for high yields occurred at 2625.7, 2548.1, and 2705.6 °C d^{-1} for ARG, ARX, and PAT, respectively. The highest growth rates for low yields occurred at 2768.5, 2312.4, and 2754.3 °C d^{-1} for ARG, ARX, and PAT, respectively (Figure 7).

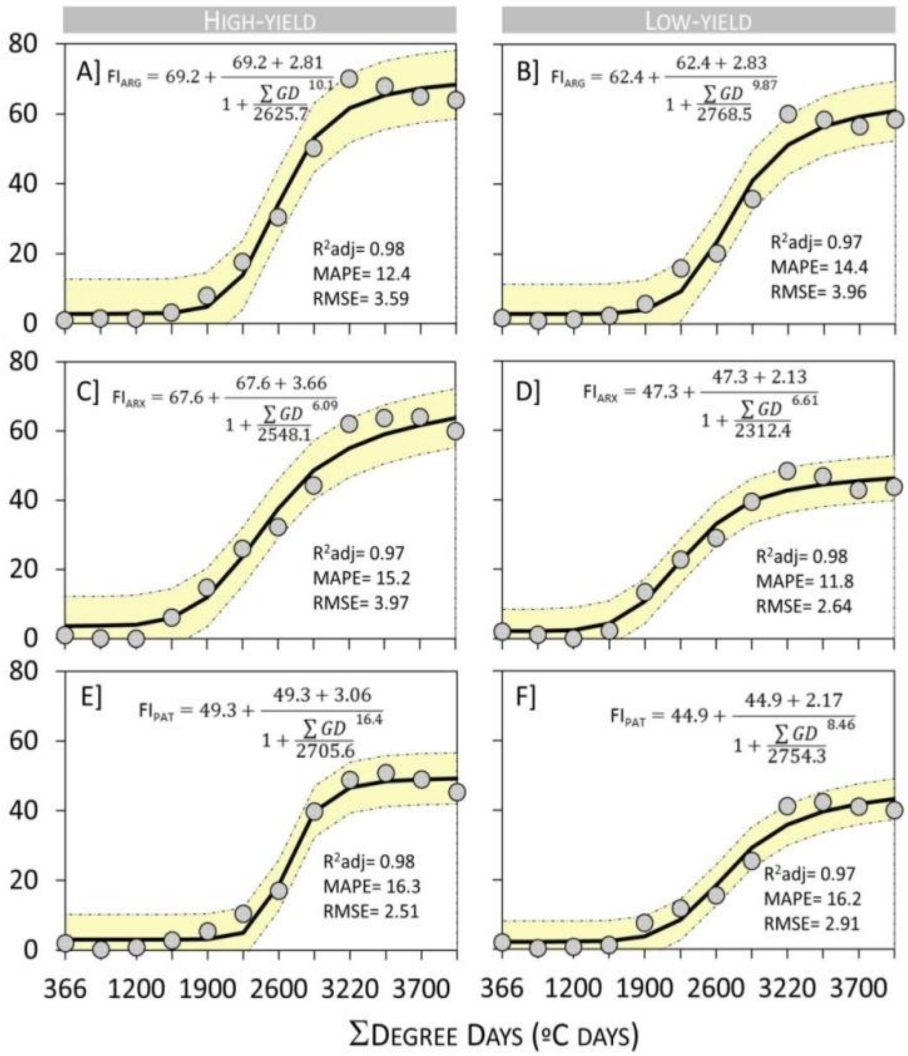

FIGURE 7. Coffee rust index (FI) based on total degree days (ΣDD, °C d^{-1}) after the beginning of vegetative growth (September) in Cerrado Mineiro. A and B, Araguari; C and D, Araxá; and E and F, Patrocínio. Points are observed values, solid lines are estimated values, and dashed lines indicate 95% confidence intervals.

The cercospora index also tended to have a sigmoidal distribution. The models indicated that the cercospora index could be estimated as a function of DD, because the adjustments were all significant (*P*<0.001), with low MAPEs and RMSEs and R^2adj ≈1.0. For example, the logistic model for CDM at low yields had a MAPE of 7.74%, RMSE of 0.40%, and R^2adj = 0.79 (Figure 8D). A MAPE of 7.74% is low, because the highest cercospora index of 3.97% at the end of the cycle had a variation of only ±0.28%.

The most accurately fitted models for estimating cercospora in SO_{MG} for high yields were $\Sigma CI_{VAR} = F(\Sigma DD)$ and $\Sigma CI_{MUZ} = F(\Sigma DD)$, with MAPEs of 3.61 and 7.69%, respectively. The most accurate models for low yields were $\Sigma CI_{CDM} = F(\Sigma DD)$ and $\Sigma CI_{MUZ} = F(\Sigma DD)$, with MAPEs of 7.74 and 7.69%, respectively.

Ymax of the adjusted model indicates the maximum disease intensity. The highest cercospora intensities for high yields were 6.29, 6.48, 6.18, and 8.07% in BOE, CDM, MUZ, and VAR, respectively (Figure 7). The highest cercospora intensities for low yields were 4.54, 3.97, 6.21, and 3.38% in BOE, CDM, MUZ, and VAR, respectively (Figure 8). Cercospora intensity in SO_{MG} occurred at an accumulation of 3654.45 °C d^{-1}. Custodio *et al.* (2010) reported maximum cercospora intensity in SO_{MG} of 20%.

The cercospora index for SO_{MG} indicated high maximum cercospora growth rates of 70.8, 2.70, 64.75, and 3.26% for high yields and 65.2, 48.3, 64.7, and 4.19% for low yields in BOE, CDM, MUZ, and VAR, respectively (Figure 8). The highest cercospora growth rate was 70.8%. This development rate occurred in a high-yield tree in BOE where the plant accumulated 1001.9 °C d^{-1} (Figure 8B). The highest rates of cercospora development in SO_{MG} for high yields occurred at 1001.9, 1391.2, 999.3, and 2227.3 °C d^{-1} for BOE, CDM, MUZ, and VAR, respectively (Figure 8A, C, E, G). The highest development rates of the disease for low yields occurred at 977.4, 1365.5, 1004.6, and 1426.2 °C d^{-1} for BOE, CDM, MUZ, and VAR, respectively (Figure 8B, D, F, H).

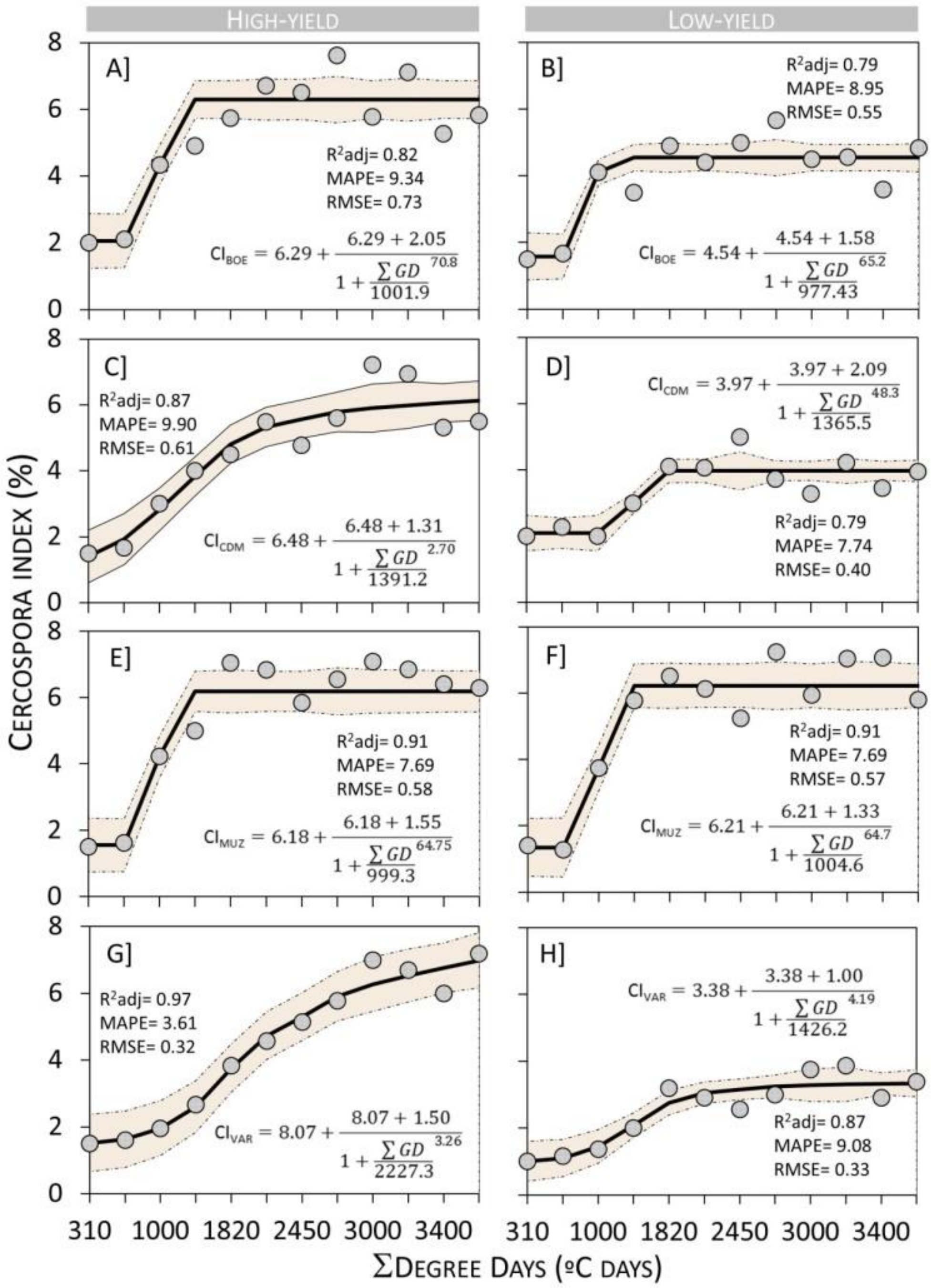

FIGURE 8. Cercospora index (CI) based on total degree days (ΣDD, °C d^{-1}) after the beginning of vegetative growth (September) in Sul de Minas Gerais. A and B, Boa Esperança; C and D, Carmo de Minas; E and F Muzambinho; G and H Varginha. Points are observed values, solid lines are estimated values, and dashed lines indicate 95% confidence intervals.

The most accurate model for estimating cercospora for high yields in CE_{MG} was $\Sigma CI_{ARG} = F(\Sigma DD)$, with a MAPE of 12.7% (Figure 9A). The most accurate model for low yields was $\Sigma CI_{PAT} = F(\Sigma DD)$, with a MAPE of 9.05% (Figure 9F). Ymax of the adjusted model indicates the highest disease intensity. The highest cercospora intensities in ARG, ARX, and PAT were 67.9, 13.1, and 12.09% for high yields (Figure 9A, C, E) and 73.6, 12.5, and 6.76% for low yields, respectively (Figure 9B, D, F). These maximum cercospora intensities in CE_{MG} occurred at an accumulation of 3994.16 °C d^{-1}.

The cercospora index in ARG, ARX, and PAT in CE_{MG} had maximum growth rates of 9.97, 60.2, and 2.01% for high yields and 5.75, 48.7, and 13.5% for low yields, respectively (Figure 9). The highest cercospora growth rate was 60.2%. This growth rate occurred in a high-yield tree in ARX where the plant accumulated 1485.2 °C d^{-1} (Figure 9B). The highest cercospora growth rates at ARG, ARX and PAT in CE_{MG} occurred at 2581.3, 1485.2 and 1553.6 °C d^{-1} for high yields (Figure 9A, C, E) and 2888.1, 1444.4 and 1161.5 °C d^{-1} for low yields (Figure 9B, D, F), respectively.

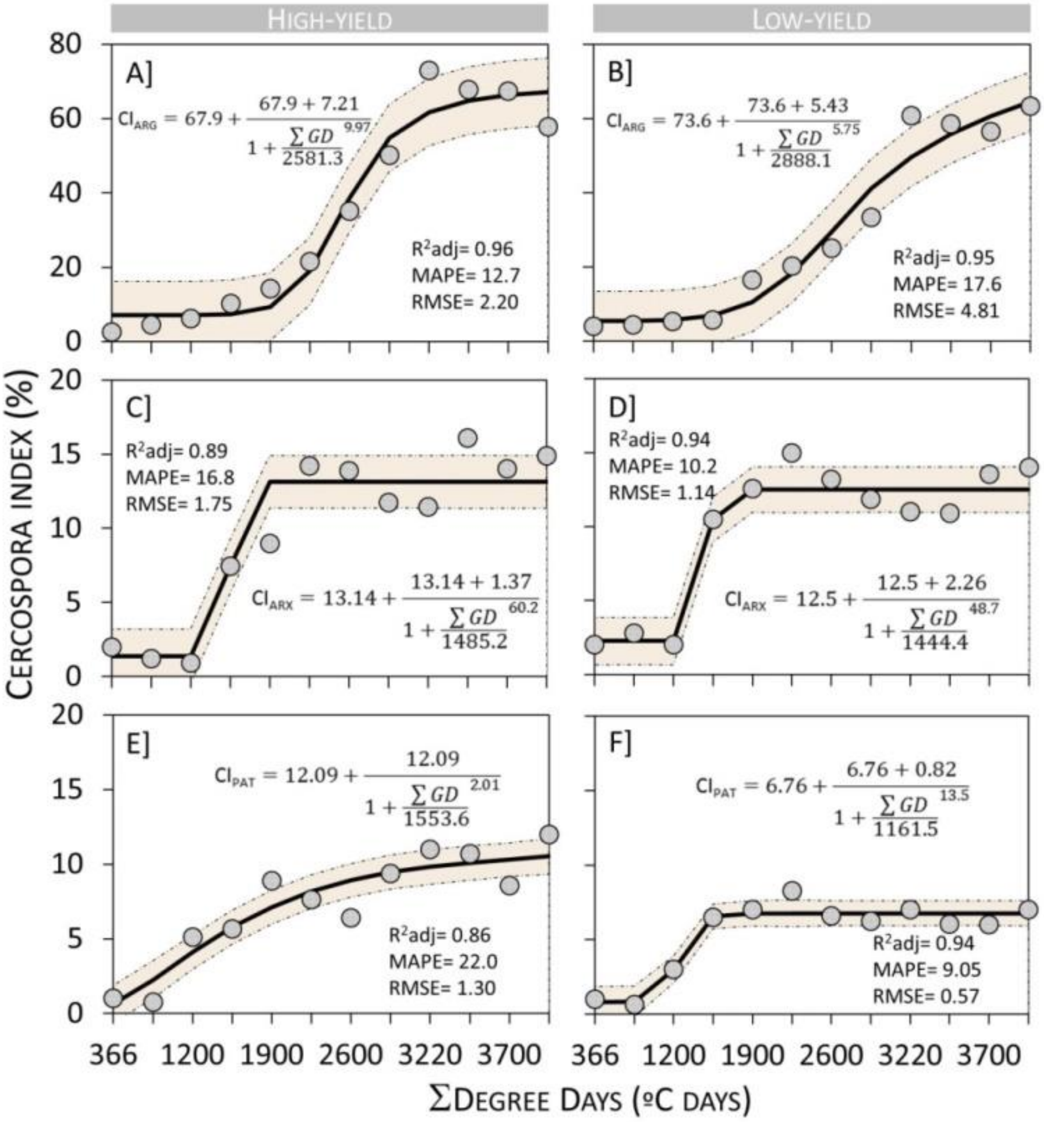

FIGURE 9. Cercospora index (CI) based on total degree days (ΣDD, °C d^{-1}) after the beginning of vegetative growth (September) in Cerrado Mineiro. A and B, Araguari; C and D, Araxá; and E and F, Patrocínio. Points are observed values, solid lines are estimated values, and dashed lines indicate 95% confidence intervals.

The severity of infestation with the berry borer tended to have a Gaussian distribution. The adjustments of the models indicated that the severity of infestation could be estimated as a function of the climatic conditions, because the adjustments were all significant (P<0.001), with a low MAPE and RMSE and R^2adj ≈1.0. The Lorentz model estimated the severity of infestation for high yields in BOE with high accuracy, with a MAPE of 8.99%, RMSE of 0.11%, and R^2adj of 0.98 (Figure 10A). A

MAPE of 8.99% was relatively low, because the average infestation was 2.5% and the maximum peak intensity had a variation of only ±0.22%.

The most accurately fitted models for estimating the severity of infestation with the berry borer for high yields in SO_{MG} were $\Sigma BC_{BOE} = F(\Sigma DD)$ and $\Sigma BC_{CDM} = F(\Sigma DD)$, with MAPEs of 8.99 and 18.10%, respectively. The most accurately fitted models with crop adjustments for low yields were $\Sigma BC_{MUZ} = F(\Sigma DD)$ and $\Sigma BC_{VAR} = F(\Sigma DD)$, with MAPEs of 18.2 and 18.6%, respectively.

The maximum intensities of infestation with the berry borer in BOE, CDM, MUZ, and VAR were 2.6, 1.9, 2.2, and 2.01% for high yields and 1.1, 1.8, 2.3, and 0.7% for low yields, respectively (Figure 10). These maximum intensities in SO_{MG} occurred at distinct accumulations for each site. For example, the pest peaks in BOE, CDM, MUZ, and VAR occurred at 2007.9, 2346.8, 2298.4, and 1995.2 °C d^{-1} for high yields (Figure 10A, C, E, G) and 2011.1, 2309, 2317.8, and 2019.3 °C d^{-1} for low yields (Figure 10B, D, F, H), respectively. A more complex analysis of the severity of infestation with the berry borer in BOE indicated an increase in severity up to 1520 °C d^{-1}. The severity then stabilised slightly and decreased after the accumulation of 2010 °C d^{-1} (Figure 10A).

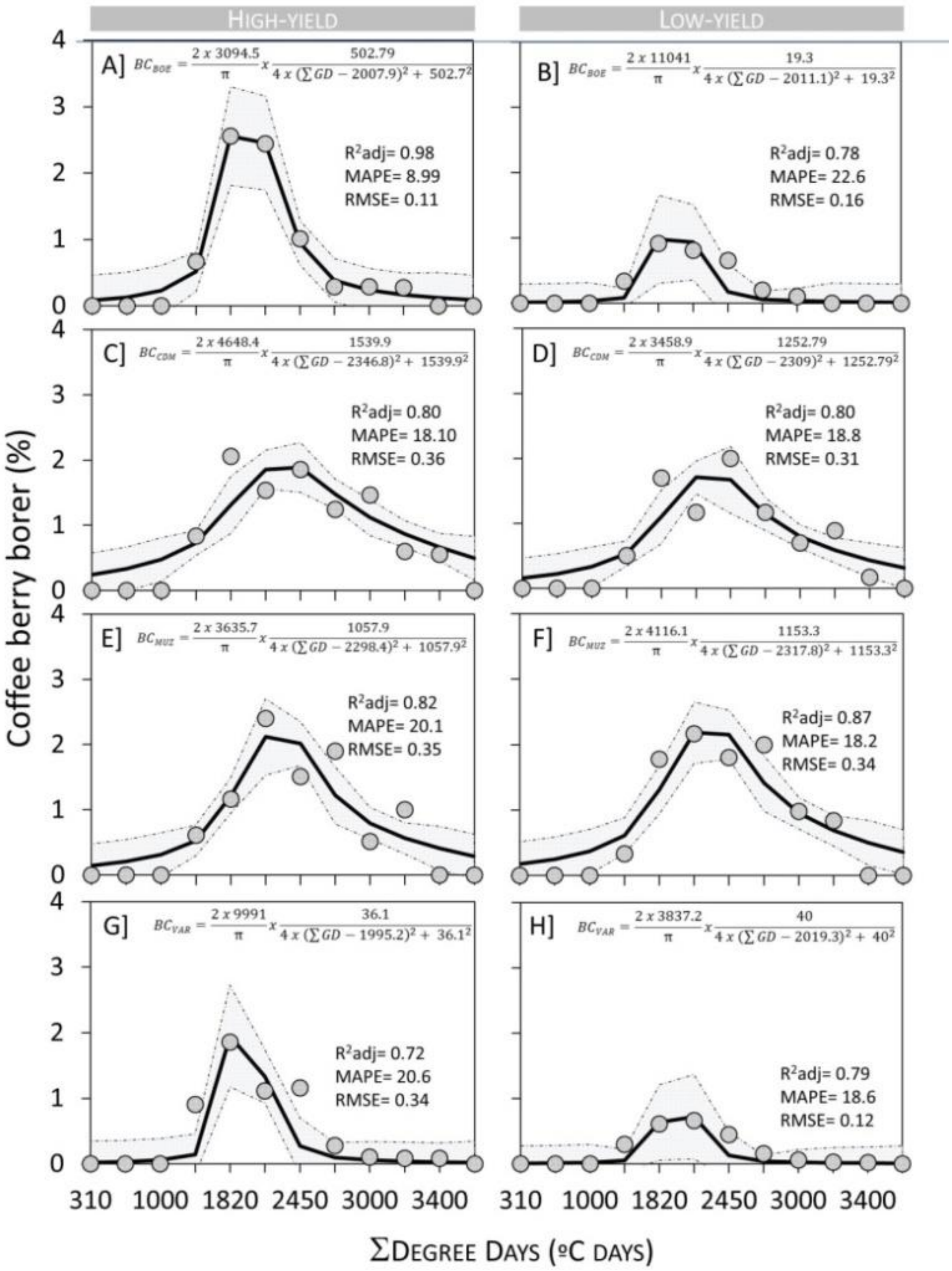

FIGURE 10. Coffee berry borer infestation based on total degree days (ΣDD, °C d^{-1}) after the beginning of vegetative growth (September) in Sul de Minas Gerais. A and B, Boa Esperança; C and D, Carmo de Minas; E and F, Muzambinho; and G and H, Varginha. Points are observed values, solid lines are estimated values, and dashed lines indicate 95% confidence intervals.

The most accurate model for estimating the severity of infestation in CE_{MG} for high yields was ΣBCARG = F(ΣDD), with a MAPE of 20.4%, RMSE of 2.36, and R^2adj = 0.81 (Figure 11A). The most accurate model for low yields was ΣBCARX = F(ΣDD), with a MAPE of 10.4%, RMSE of 0.29%, and R^2adj of 0.99 (Figure 11D). The maximum intensities of infestation in ARG, ARX, and PAT in CE_{MG} were 16.2, 3.1, and 6.6% for high yields (Figure 11A, C, E) and 13.5, 11.9, and 7.1% for low yields (Figure 11B, D, F), respectively.

Pest infestation in ARG, ARX, and PAT peaked at 2716.2, 2373.2, and 2576.3 °C d^{-1} for high yields (Figure 11A, C, E, G) and 2528.1, 2408.3 and 2532.7 °C d^{-1} for low yields (Figure 11B, D, F, H), respectively.

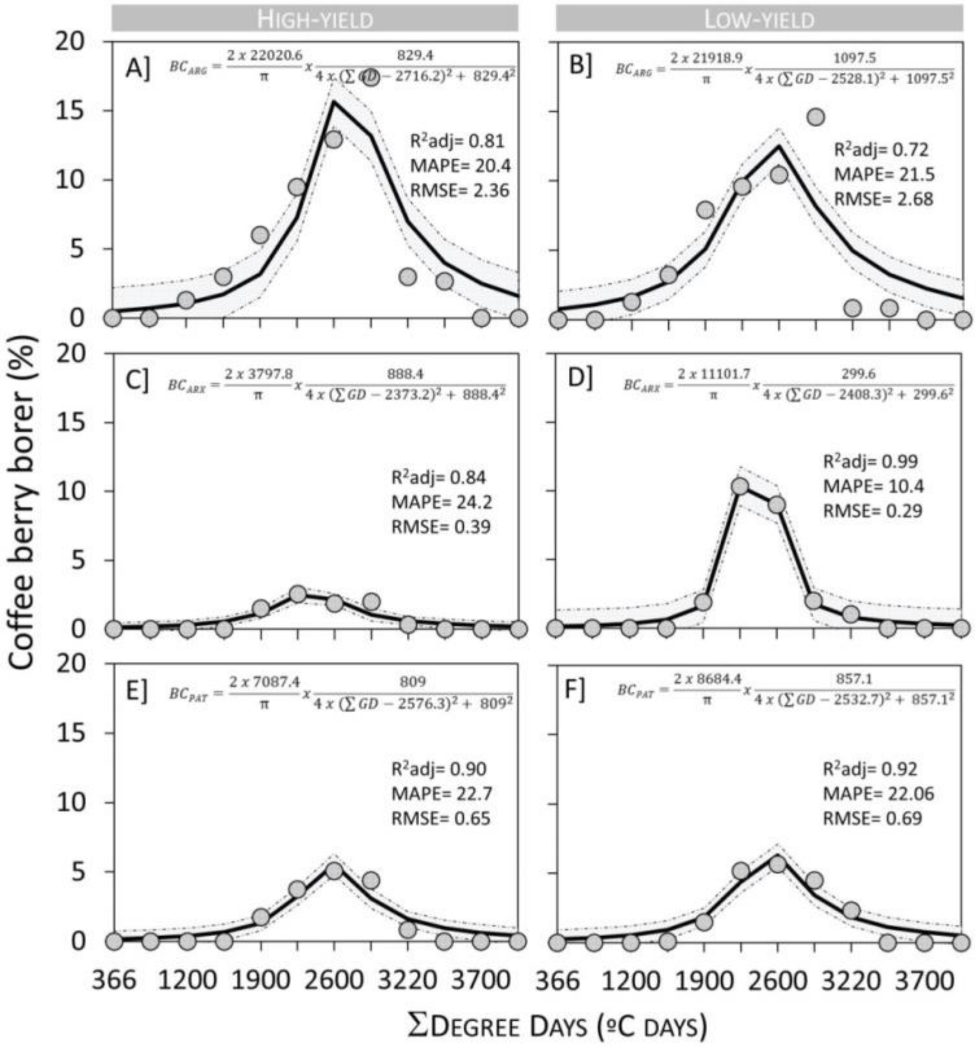

FIGURE 11. Coffee berry borer infestation based on total degree days (ΣDD, °C d^{-1}) after the beginning of vegetative growth (September) in Cerrado Mineiro. A and B, Araguari; C and D, Araxá; E and F, Patrocínio. Points are observed values, solid lines are estimated values, and dashed lines indicate 95% confidence intervals.

The severity of infestation with the leaf miner tended to have an inverted peak for all SO_{MG} sites. The adjustments of the models indicated that the severity of infestation could be estimated as a function of the climatic conditions, because the adjustments were all significant (P<0.001), with low MAPE and RMSE and R^2adj ≈1.0. The Lorentz model accurately estimated the severity of infestation with leaf miner for high yields in MUZ, with a MAPE of 5.80%, RMSE of 0.40%, and R^2adj of 0.74 (Figure 12E). An error of 5.80% is low, because a severity of 2.2% has a variation of only ±0.12%. The most accurate models for estimating the severity of infestation were $\Sigma BM_{MUZ} = F(\Sigma DD)$ and $\Sigma BM_{VAR} = F(\Sigma DD)$, with MAPEs of 5.80 and 5.97%, for high yields and $\Sigma BM_{MUZ} = F(\Sigma DD)$ and $\Sigma BM_{CDM} = F(\Sigma DD)$, with MAPEs of 5.22 and 10.1%, for low yields, respectively.

The level of infestation was assessed after the coffee harvest and when new shoots appeared. Infestation with the leaf miner was already high at the beginning of the evaluations, represented by *Y*min of the equation of Lorentz. Infestations at the beginning of the evaluations in BOE, CDM, MUZ, and VAR were 2.74, 2.06, 2.60, and 1.48% for high yields and 3.04, 2.08, 2.66 and 1.70% for low yields, respectively (Figure 12). The lower intensities of the leaf miner in SO_{MG} occurred at 1000-2400 °C d^{-1}. High rainfall during this period may account for the low intensities (Figure 2). Infestation was higher in other periods when rainfall was lower. Several studies such as Villacorta (1980) and Custódio *et al.* (2009) verified that the population density of the leaf miner was highest after periods of P.

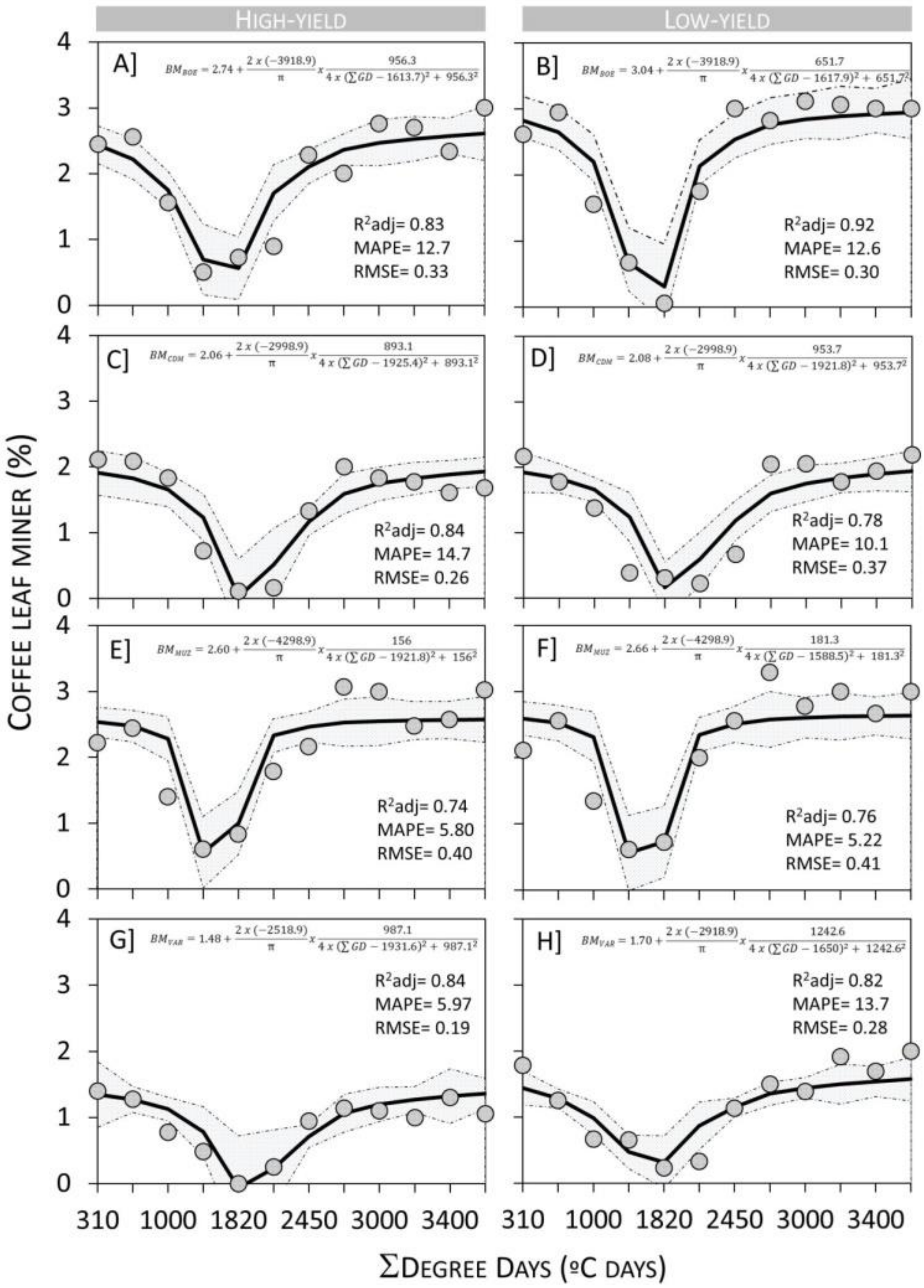

FIGURE 12. Coffee leaf miner infestation based on total degree days (ΣDD, °C d^{-1}) after the beginning of vegetative growth (September) in Sul de Minas Gerais. A and B, Boa Esperança; C and D, Carmo de Minas; E and F, Muzambinho; and G and H, Varginha. Points are observed values, solid lines are estimated values, and dashed lines indicate 95% confidence intervals.

The severity of infestation with the leaf miner had distinct trends amongst the

CE_{MG} sites. For example, infestation in ARX tended to have an inverted peak (vale curve), but the other sites had sigmoid curves. Regardless of the trend, adjustments of the models indicated that the severity of infestation could be estimated as a function of the climatic conditions, because the adjustments were all significant (P<0.001), with a low MAPE and RMSE and R^2adj ≈1.0. Infestation of leaf miner was high in ARG and PAT. The highest rates in ARG and PAT were 42.3 and 20.2% for high yields and 40.5 and 21.9% for low yields, respectively (Figure 13).

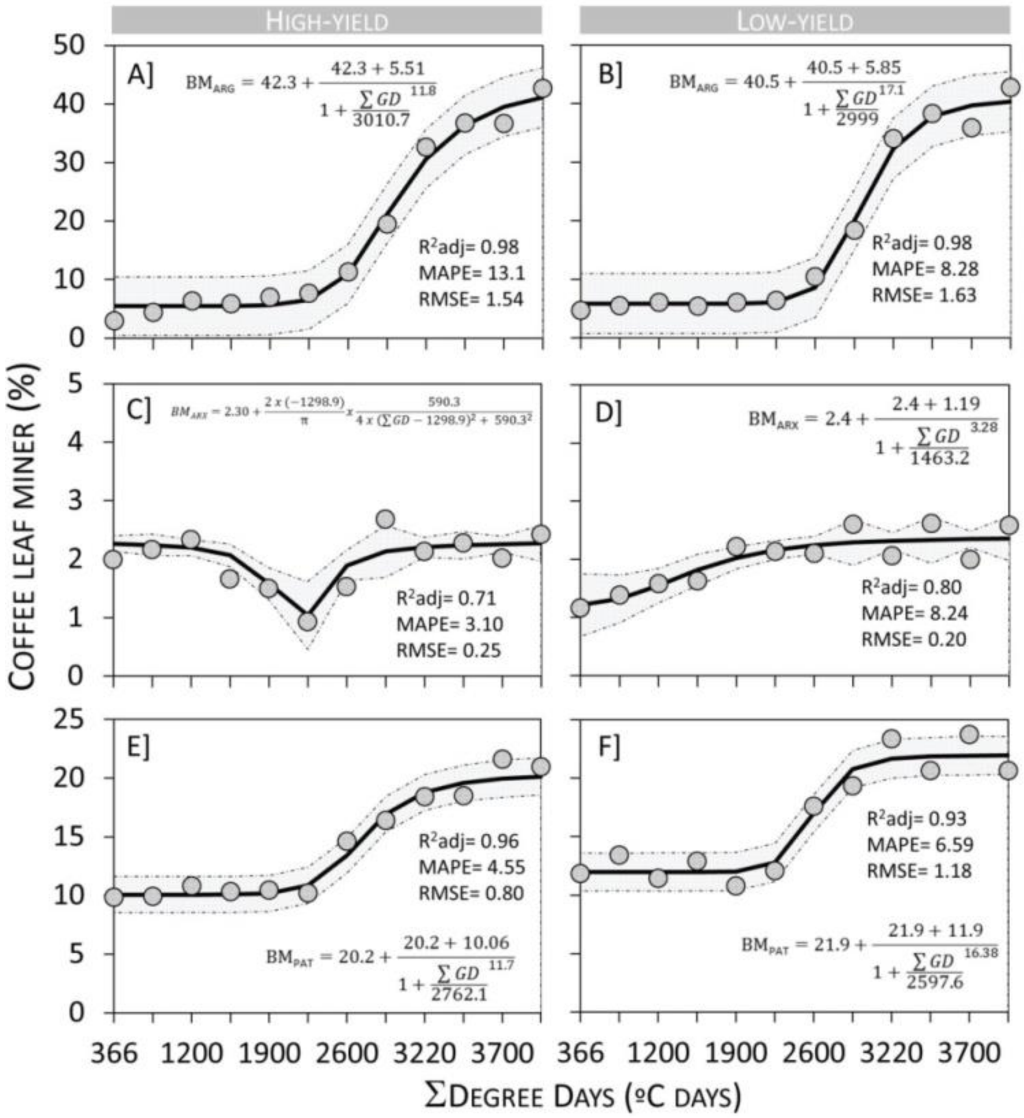

FIGURE 13. Coffee leaf miner infestation based on total degree days (ΣDD, °C d^{-1}) after the beginning of vegetative growth (September) in Cerrado Mineiro. A and B, Araguari; C and D, Araxá; and E and F, Patrocínio. Points are observed values, solid lines are estimated values, and dashed lines indicate 95% confidence intervals.

Summaries of the adjusted parameters in the trend models for all sites are presented in Tables 4-6.

TABLE 4. Parameters adjusted in the models for the number of nodes and leafiness as functions of total Degree Days.

REGION	City	Ymax	Ymin	X_0	*p*	Ymax	Ymin	X_0	*p*
		Number of nodes				Leafiness			
South	BOE	10.09	1.22	1744.6	1.87	98.69	43.26	3041	12.34
	CDM	8.47	1.22	1533.5	2.16	99.1	18.37	3234.8	7.05
	MUZ	11.9	1.18	2250.7	1.66	95.98	51.5	2829.7	13.1
	VAR	8.16	1.24	1498.7	2.23	99.02	34.44	3009.8	13.14
Cerrado	ARG	27.1	0.45	6406.3	1.11	96.37	0	3557.8	6.58
	ARX	11.15	2.21	2351.4	2.3	96.44	17.19	3250.4	6.48
	PAT	10.6	1.71	2136.7	2.07	97.49	22.29	3509.7	9.85
Average (±Standard deviation)	South	9.66 (±1.719)	1.22 (±0.025)	1756.88 (±346.682)	1.98 (±0.264)	98.2 (±1.489)	36.89 (±14.178)	3028.83 (±165.918)	11.41 (±2.928)
	Cerrado	16.28 (±9.372)	1.46 (±0.907)	3631.47 (±2405.473)	1.83 (±0.631)	96.77 (±0.627)	13.16 (±11.679)	3439.3 (±165.351)	7.64 (±1.917)
	Total	12.5 (±6.581)	1.32 (±0.54)	2560.27 (±1729.996)	1.91 (±0.418)	97.58 (±1.351)	26.72 (±17.518)	3204.74 (±266.492)	9.79 (±3.094)

TABLE 5. Parameters adjusted in the models for the coffee leaf rust and cercospora indices as functions of total Degree Days.

REGION	YIELD	City	Ymax	Ymin	X_0	*p*	Ymax	Ymin	X_0	*p*
			Rust index				Cercospora index			
South	HIGH	BOE	72.3	3.05	2542	5.6	6.29	2.05	1002	70.8
	HIGH	CDM	81.4	3.32	2418	4.36	6.48	1.31	1391	2.7
	HIGH	MUZ	84.9	0.74	2792	3.25	6.18	1.55	999	64.8
	HIGH	VAR	81	4.01	2215	6.08	8.07	1.5	2227	3.26
	LOW	BOE	18.5	1.97	2485	6.75	4.54	1.58	977	65.2
	LOW	CDM	35.3	2.09	2264	3.54	3.97	2.09	1366	48.3
	LOW	MUZ	34.4	1.87	2211	3.82	6.21	1.33	1005	64.7
	LOW	VAR	45.2	2.07	2246	3.41	3.38	1	1426	4.19
Cerrado	HIGH	ARG	69.2	2.81	2626	10.1	67.9	7.21	2581	9.97
	HIGH	ARX	67.6	3.66	2548	6.09	13.1	1.37	1485	60.2
	HIGH	PAT	49.3	3.06	2706	16.4	12.1	0	1554	2.01
	LOW	ARG	62.4	2.83	2769	9.87	73.6	5.43	2888	5.75
	LOW	ARX	47.3	2.13	2312	6.61	12.5	2.26	1444	48.7
	LOW	PAT	44.9	2.17	2754	8.46	6.76	0.82	1162	13.5
Average (±Standard deviation)	South-High		79.89 (±5.373)	2.78 (±1.419)	2491.75 (±241.184)	4.82 (±1.274)	6.76 (±0.885)	1.6 (±0.316)	1404.93 (±578.345)	35.38 (±37.492)
	South-Low		33.35 (±11.031)	2 (±0.101)	2301.23 (±124.14)	4.38 (±1.589)	4.53 (±1.219)	1.5 (±0.459)	1193.43 (±235.312)	45.6 (±28.7)
	Cerrado-High		62.03 (±11.056)	3.18 (±0.437)	2626.47 (±78.753)	10.86 (±5.197)	31.04 (±31.923)	2.86 (±3.829)	1873.37 (±614.041)	24.06 (±31.55)
	Cerrado-Low		51.53 (±9.487)	2.38 (±0.393)	2611.73 (±259.327)	8.31 (±1.635)	30.95 (±37.044)	2.84 (±2.358)	1831.33 (±926.053)	22.65 (±22.89)
	South		56.62 (±26.14)	2.39 (±1.02)	2396.49 (±204.709)	4.6 (±1.354)	5.64 (±1.547)	1.55 (±0.369)	1299.18 (±424.101)	40.49 (±31.389)
	Cerrado		56.78 (±10.862)	2.78 (±0.575)	2619.1 (±171.599)	9.59 (±3.718)	31 (±30.928)	2.85 (±2.844)	1852.35 (±703.121)	23.36 (±24.665)

TABLE 6. Parameters adjusted in the models for infestations with the coffee leaf miner and coffee berry borer as functions of total Degree Days.

REGION	YIELD	City	Ymin	A	X_0	W	Ymin	A	X_0	W
			Leaf miner				Berry Borer			
South	HIGH	BOE	2.74	-3919	1614	956	0	3095	2008	503
	HIGH	CDM	2.06	-2999	1925	893	0	4648	2347	1540
	HIGH	MUZ	2.6	-4299	1922	156	0	3636	2298	1058
	HIGH	VAR	1.48	-2519	1932	987	0	9991	1995	36.1
	LOW	BOE	3.04	-3919	1618	652	0	11041	2011	19.3
	LOW	CDM	2.08	-2999	1922	954	0	3459	2309	1253
	LOW	MUZ	2.66	-4299	1589	181	0	4116	2318	1153
	LOW	VAR	1.7	-2919	1650	12426	0	3837	2019	40
Cerrado	HIGH	ARX	2.3	-1299	1299	590				
			Ymax	Ymin	X_0	p	Ymin	A	X_0	W
	HIGH	ARG	42.3	5.51	3011	11.8	0	22020.6	2716	829
	HIGH	ARX	-	-	-	-	0	3798	2373	888
	HIGH	PAT	20.2	10.1	2762	11.7	0	7087	2576	809
	LOW	ARG	40.5	5.85	2999	17.1	0	21918.9	2528	1098
	LOW	ARX	2.4	1.19	1463	3.28	0	11101.7	2408	300
	LOW	PAT	21.9	11.9	2598	16.4	0	8684	2533	857
Average (±Standard deviation)	South-High		2.22 (±0.574)	-3433.9 (±818.515)	1848.13 (±156.336)	748.13 (±396.684)	0 (±0)	5342.4 (±3165.28)	2162.08 (±186.481)	784.17 (±654.431)
	South-Low		2.37 (±0.596)	-3533.9 (±682.618)	1694.55 (±153.568)	3553.18 (±5923.749)	0 (±0)	5613.3 (±3628.476)	2164.3 (±172.236)	616.35 (±678.729)
	Cerrado-High		31.25 (±15.627)	7.79 (±2.535)	2886.4 (±140.274)	11.75 (±3.089)	0 (±0)	10968.6 (±9711.611)	2555.23 (±172.468)	842.27 (±41.234)
	Cerrado-Low		21.6 (±19.052)	6.31 (±7.573)	2353.27 (±802.142)	12.25 (±9.263)	0 (±0)	13901.67 (±7047.543)	2489.7 (±70.532)	751.4 (±409.317)
	South		2.3 (±0.548)	-3483.9 (±699.775)	1771.34 (±165.289)	2150.65 (±4165.86)	0 (±0)	5477.85 (±3155.523)	2163.19 (±166.189)	700.26 (±623.72)
	Cerrado		25.46 (±16.446)	6.9 (±4.202)	2566.52 (±640.482)	12.05 (±5.508)	0 (±0)	12435.13 (±7757.201)	2522.47 (±123.192)	796.83 (±264.902)

The most accurate trend models were for estimating leaf miner infestation, and the least accurate models were for estimating berry borer infestation, with MAPEs averaging 8.90, 19.13%, respectively (Table 7). The accuracy of the models was similar for both high and low yields, with MAPEs of 12.44 and 13.00% for SO_{MG} and 15.31 and 13.04% for CE_{MG}, respectively.

TABLE 7. Synthesis of the statistical indices (MAPE, R^2adj, and RMSE) by region used for comparing the observed data and the adjusted nonlinear models for the diseases and severity of the pest infestations for Sul de Minas Gerais and Cerrado Mineiro.

	South of Minas		Cerrado Mineiro		
	High yield	Low yield	High yield	Low yield	Total
Diseases and pests	MAPE				
Coffee leaf miner	9,775	10,405	6,941	7,708	8,905
Coffee berry borer	16,946	19,660	22,494	17,995	19,135
Cercospora index	7,640	8,368	17,189	12,345	10,902
Rust index	15,403	13,585	14,633	14,133	14,446
Average	12,441	13,004	15,314	13,045	13,347
standard deviation	4,443	4,929	6,471	4,268	4,488
	R^2_{adj}				
Coffee leaf miner	0,818	0,820	0,891	0,908	0,854
Coffee berry borer	0,836	0,813	0,856	0,881	0,843
Cercospora index	0,898	0,846	0,905	0,946	0,895
Rust index	0,981	0,974	0,983	0,975	0,978
Average	0,883	0,863	0,909	0,928	0,893
standard deviation	0,074	0,075	0,053	0,041	0,061
	RMSE				
Coffee leaf miner	0,297	0,360	0,868	1,009	0,590
Coffee berry borer	0,294	0,238	1,141	1,225	0,659
Cercospora index	0,563	0,468	2,752	2,177	1,351
Rust index	3,425	1,604	3,362	3,175	2,838
Average	1,145	0,667	2,031	1,897	1,359
Standard deviation	1,525	0,632	1,216	0,992	1,044

We applied all possible combinations for estimating pest infestation and disease severity in SO_{MG} and CE_{MG}. The estimation models were calibrated as a function of the coffee leafiness estimated by DD (LF = F(ΣDD)), the growth of nodes of plagiotropic branches estimated by DD (NN = F(ΣDD)), and accumulated by ΣDD. We could thus estimate LF and NN with the trend models using only DD, and the results of these models could then be used to estimate the rates of pest infestation and disease severity. For example, %CER = f[ΣDD, LF = f(ΣDD), and NN = f(ΣDD)]. All models were accurate, with MAPEs ranging from 5.6567% (coffee rust in SO_{MG}) to 19.587% (cercospora in SO_{MG}) and all R^2adj were >0.4229 (Table 8). The performance of all estimation models is shown in Figure 14.

TABLE 8. Regional models for simulating coffee diseases and pests as functions of leafiness, number of nodes, and degree days accumulated from the beginning of vegetative growth (September 01). CL, linear coefficient; NN, number of nodes estimated by ΣDD; LE, leafiness estimated by ΣDD; and ΣDD, degree days accumulated from 1 September.

Diseases and pests	Regions	Coefficients				Indexes			
		CL	NN	LEA	ΣDD	p-valor	MAPE	RMSE	R^2adj
Rust index	South	22.72	7.98	-0.358	-0.00568	0.0001	19.58	4.425	0.933
	Cerrado	22.72	7.98	-0.358	-0.00568	0.0001	17.64	4.200	0.965
Cercospora	South	4.322	1.36	-0.031	-0.00213	0.0001	5.656	0.431	0.726
	Cerrado	28.47	5.76	-0.341	-0.00919	0.0001	7.690	1.300	0.983
Coffee leaf miner	South	3.854	-0.459	-0.017	0.00091	0.0001	9.472	0.687	0.422
	Cerrado	27.21	-0.073	-0.213	0.00035	0.0001	6.056	1.433	0.927
Coffee berry borer	South	-4.06	0.245	0.034	0.00026	0.0001	15.82	0.252	0.814
	Cerrado	-26.56	2.801	0.202	-0.00138	0.0001	14.81	1.685	0.714

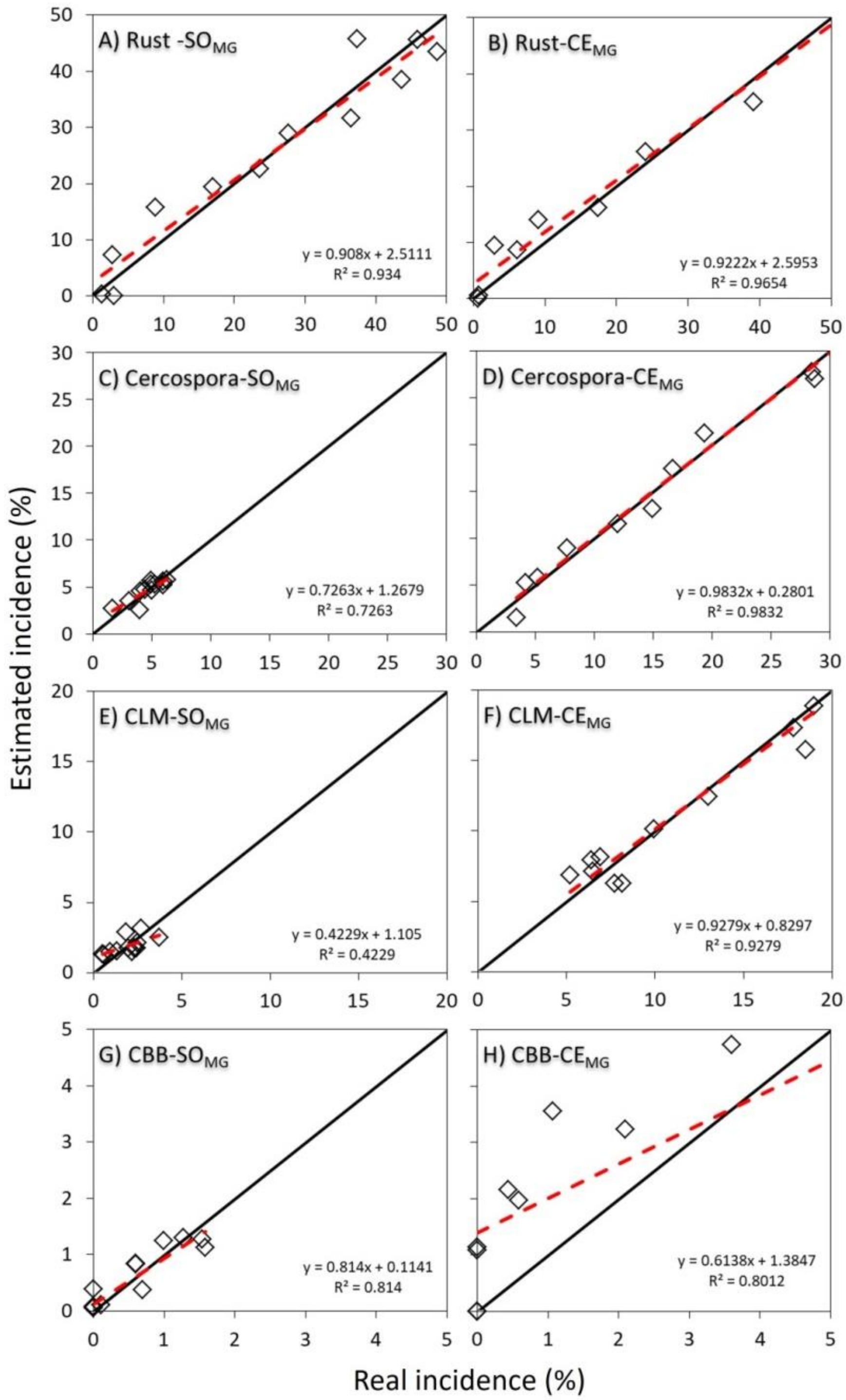

FIGURE 14. Relationships between model estimates and observed incidences of coffee rust, cercospora, coffee leaf miner (CLM), and coffee berry borer (CBB) in Sul de Minas (SO_{MG}) and Cerrado Mineiro (CE_{MG}). The dashed lines are trend lines and the solid lines are 1:1 lines.

The model for estimating the severity of cercospora in SO_{MG} was accurate (Eq. 12), with a MAPE, RMSE, and R^2adj of 5.656%, 0.431%, and 0.726, respectively, and was significant at *P*<99% (Table 8). A MAPE of 5.656% is considered low in studies modelling crops (Martins *et al.,* 2015; Moreto *et al.,* 2015), e.g. in MUZ where rust severity was 34.4% for low yields, with an error of only ±1.943%.

$$\%\mathrm{CER}_{SO} = -0.00213 \cdot \mathrm{DD} - 0.0314 \cdot \mathrm{LFf}(\sum \mathrm{DD}) + 1.369 \cdot \mathrm{NNf}(\sum \mathrm{DD}) + 4.322 \quad (12)$$

NN was the most sensitive variable calibrated in the models and was positively correlated with cercospora severity. The result of the model indicated that the severity of disease increased with plant density. Every 10 units of NN in this equation represent an increase of 13.69% in the severity of cercospora in the coffee crop.

Conclusions

Coffee planted in CE_{MG} had higher rates of diseases and pest infestation than the coffee planted in SO_{MG}, due to higher air temperatures. The levels of pest infestation and disease severity were highest in ARG in CE_{MG}. Trees at this site with high yields had rust, cercospora, leaf miner, and berry borer intensities of 30.9, 36.1, 18.82, and 4.5%, respectively.

The intensities of the rust index varied on the yields of the coffee trees. Cercospora and pest intensities were independent of yield. The trends of pest infestation and disease severity over time could be estimated as functions of the thermal index. Rust and cercospora severity tended to have sigmoidal distributions over time, whereas leaf miner and berry borer infestations tended to have Gaussian distributions. The trend models of pest infestation and disease severity as functions of degree days were accurate, with low error (MAPE) and high accuracy (R^2adj » 1.0). FI_{MUZ} = f(DD) for low yields and BC_{BOE} = f(DD) for high yields were the most accurate models for SO_{MG}, and FI_{ARX} = f(DD) for low yields and CI_{PAT} = f(DD) for low yields were the most accurate models for CE_{MG}.

The models for estimating pest infestation and disease severity as functions of leafiness, NN, and DD were accurate, with MAPEs ranging from 5.656% (rust for SO_{MG}) to 19.587% (cercospora for SO_{MG}). Leafiness (LF = F(ΣDD)) and the of nodes

($NN = F(\Sigma DD)$) were estimated by the trend models using DDs (using only air temperature), indicating that a phytosanitary warning system could be established and that the levels of pest infestation and disease severity of coffee could be estimated.

Acknowledgements

This research was supported by the São Paulo Research Foundation (FAPESP, grant #: 2015/17797-4).

CAPÍTULO 3 – Machine learning algorithms for forecasting the incidence of Coffea arabica pests and diseases

ABSTRACT - The production and quality of coffee in Brazil are affected by phytosanitary problems: coffee rust (*Hemileia vastatrix*), cercospora (*Cercospora coffeicola*), coffee miner (*Leucoptera coffeella*) and coffee borer (*Hypothenemus hampei*). The intensity of these phytosanitary problems is controlled by climatic variability, and few studies address the problem due to their complexity. Disease and pest alert models are able to generate information for agrochemical applications only when needed, reducing costs and environmental impacts. With machine learning algorithms it is possible to develop models to be used in disease and pest warning systems as a function of the climate in order to improve the efficiency of chemical control of pests of the coffee tree. Thus, we evaluated progress curves of the incidence of pests and diseases, correlated the infection rates with the meteorological variables, and also calibrated and tested machine learning algorithms to predict the incidence of coffee rust, cercospora, coffee miner and, coffee borer. We used climatic and field data obtained from coffee plantations in production in the southern regions of the State of Minas Gerais and from the region of the Cerrado Mineiro; these crops did not receive phytosanitary treatments. As dependent variables, we considered the monthly rates of coffee rust, cercospora, coffee miner and coffee tree borer, and the climatic elements were considered as independent (predictor) variables. All analyses were performed considering three different time periods: 1-10d (from 1 to 10 days before the incidence evaluation); 11-20d (from 11 to 20 days before the incidence assessment); and 21-30d (from 21 to 30 days before the incidence assessment). The Pearson coefficient was used to evaluate the unit correlations between the meteorological variables and infection rates of coffee diseases and pests. The algorithms calibrated and tested for prediction were: a) Multiple linear regression (RLM); b) K Neighbors Regressor (KNN); c) Random Forest Regressor (RFT) and d) Artificial Neural Networks (MLP). The models were calibrated in years of high and low-yields, because the biannual variation of harvest yield of coffee beans influences the severity of the diseases. The models were compared by the Willmott's 'd', RMSE (root mean square error) and coefficient of determination (R2) indices. The result of the more accurate algorithm was specialized for the south of Minas Gerais and Cerrado Mineiro regions using the kriging method. The climatic variables that showed significant correlations with coffee rust disease were maximum air temperature, number of days with relative humidity above 80% and relative humidity. Random Forest was more accurate in the prediction of coffee rust, cercospora, coffee miner and coffee borer using climatic conditions. In the southern region of Minas Gerais, Random Forest showed a greater accuracy in the predictions for the Cerrado Mineiro in years of high and low-yields and for all diseases.

KEY-WORDS: crop modeling; big data; artificial intelligence; phytosanitary maps

Introduction

Coffee is the second most important agricultural commodity in the world, with annual revenues of around $ 24 billion (FAO, 2015, Kouadio *et al.,* 2018). The harvest of 2018/2019 was about 168 million bags of 60 kg (ICO, 2019). World coffee production is dominated by Coffea arabica, with a 64.5% share, being produced in more than 60 countries (ICO, 2017a, ICO, 2017b, Geeraert *et al.,* 2019). Developing countries are the main coffee producers (Hinnah *et al.,* 2018). Brazil led world production with 61 million bags in the 2018 harvest (ICO, 2019), with the State of Minas Gerais accounting for approximately 50% of Brazilian coffee production (CONAB, 2019).

One of the main constraints to coffee production is the damage caused by insect pests and plant diseases (Alba-Alejandre *et al.,* 2018). The plant is very susceptible to many diseases and pests and can be affected by several of these simultaneously; however, studies have focused on studies of diseases and pests in coffee individually (Avelino *et al.,* 2018). Leaf pests and diseases can be severe in perennial crops (Barbosa *et al.,* 2007, Cerda *et al.,* 2017), and coffee productivity losses reported in Brazil of 13% to 45% in 2017 (Cerda *et al.,* 2017).

The production and quality of coffee are heavily influenced by diseases and pests, and these are dependent on climatic conditions (Verhage *et al.,* 2017; Harvey *et al.,* 2018). Thus, there is a need for the development of predictive models for the incidence of pests and diseases that can improve the interpretation of the crop cycle according to the climate, incorporating climate-soil-plant factors (Malau *et al.,* 2018; Badnakhe *et al.,* 2018).

Rust is a devastating disease in coffee production in Brazil and other countries, and regions such as Colombia, Central America, Mexico, Peru and Ecuador (Avelino *et al.,* 2015). The climatic conditions of the producing regions present excellent conditions of temperature and humidity of the air for germination and invasion of coffee rust (Hinnah *et al.,* 2018). The main symptom of this disease is the defoliation of the plant, which consequently reduces its flowering and fruiting. Coffee production can be reduced from 30% to 90% due to the presence of rust (Santana *et al.,* 2018).

Rust is favored by air temperatures ranging from 20 and 25 °C and total precipitation greater than 30 mm. However, the rust epidemic increases rapidly in temperatures between 15 and 18 °C (Pereira *et al.,* 2008). Several studies have shown that the degree of infection of coffee rust is directly related to climatic conditions (Talamini *et al.,* 2003).

Cercospora is one of the main diseases of the coffee tree in Brazil, decreasing the productivity of the crop and the quality of the beverage (Botelho *et al.,* 2017; Silva *et al.,* 2019; Chaves *et al.,* 2018), and is caused by the fungus Cercospora coffeicola (Berkeley and Cooke). Environmental factors such as water stress and nutritional deficiency of the plant may lead to the occurrence of cercosporiosis (Chaves *et al.,* 2018).

Hypothenemus hampei Ferrari (Coleoptera: Curculionidae), commonly known as 'coffee borer', is the most damaging coffee pest worldwide (Plata-Rueda *et al.,* 2019). The females of this insect perforate and oviposit inside the coffee fruit where the larvae feed (Reyes *et al.,* 2019). The damage causes significant losses in productivity and alters the flavor profile of the coffee beans (Veja *et al.,* 2009; Walker *et al.,* 2019). The coffee borer population increased 8.8% with the temperature increase of 1 °C in Uganda (Gichimu, 2013).

The coffee miner (Leucoptera coffeella) is one of the main pests in Brazil (Sabino *et al.,* 2018). Their larvae penetrate the mesophyll, causing the destruction of the parenchyma. The main symptoms are leaf blade necrosis, photosynthesis reduction, and premature leaf fall (Androcioli *et al.,* 2018). The occurrence of this pest in the plant is directly related to the physiological state and the growth characteristics of the coffee trees, and in turn, is related to management practices and mainly to the predominant climatic conditions, especially with high rainfall events (Sabino *et al.,* 2018, Caramori *et al.,* 2004; Morais *et al.,* 2007).

All the relationships between the variability of pests and diseases of the coffee tree and climatic elements can be simulated with agrometeorological models (Rolim *et al.,* 2008) using machine learning algorithms (Sahoo *et al.,* 2017). Machine learning is a method that works with data analysis and seeks to automate the construction of analytical models (Shekoofa *et al.,* 2014; Li *et al.,* 2016). It is a field of computer science that works with the recognition of patterns using the computational learning theory in artificial intelligence (Sahoo *et al.,* 2017). Huber and Gillespie (1992) report that machine learning algorithms can automate disease and pest alert systems, guiding farmers' decision-making when it is necessary to use chemical control in crops.

The modelling of pests and diseases using different techniques of Machine learning is one of the important guidelines to reduce impacts on crop productivity (Donatelli *et al.,* 2017; Badnakhe *et al.,* 2018). The use of Machine Learning to detect plant diseases, weeds, water stress, prediction and estimation of crop productivity, among other agricultural operations will be routine in agriculture in the near future (Rehman *et al.,* 2019); however, there are problems in selecting the calibration model with the best predictability (Das *et al.,* 2018).

Machine learning algorithms are very promising for faster, more efficient and accurate large-scale analyses. Some examples of machine learning algorithms are k-Nearest Neighbor, Linear Regression, Artificial Neural Networks and Random Forest (Badnakhe *et al.,* 2018).

The k-Nearest Neighbor (kNN) is a simple and easy-to-implement classifier, obtained by the detection of the nearest neighbours, which are used to determine the classes. These neighbours should be chosen carefully to obtain good results (Lu et al., 2017), and the KNN method has shown satisfactory accuracies for estimated planting date (Gümüşçü *et al.,* 2019), water stress (Ravindranath *et al.,* 2018) and disease detection (Lu *et al.,* 2017; Cao *et al.,* 2018).

Artificial neural networks (ANNs) are dynamic systems and their structure is inspired by the human brain (Krishna *et al.,* 2019). However, the main problem in implementation of ANNs is to find the ideal number of neurons or hidden nodes (Das *et al.,* 2018). ANNs are useful when data diversity is very large and the relationships and interactions are nonlinear (Pourmohammadali *et al.,* 2019), as well as in agroclimatic zoning (Aparecido *et al.,* 2018; 2019), prediction and estimation of productivity (Pourmohammadali *et al.,* 2019; Ma *et al.,* 2019), pest prediction (Cai et al., 2019), water deficit (Krishna *et al.,* 2019) and disease prediction (Wheeler *et al.,* 2019; Kim *et al.,* 2018).

The Random Forest algorithm (RFR) is a widely used set learning method that is based on the merge of decision trees as the base learner (Breiman, 2001; Xu *et* *al.,* 2019). The final forecast is the mean of the predicted values of all trees (Ji *et al.,* 2019). This method has high classification accuracy and has powerful measurement capability of the most important variable (Xu *et al.,* 2019).

Multiple linear regression (RLM) is one of the simplest methods and a standard approach for model development and is widely used in agriculture (Correia *et al.,* 2015, Carvalho Júnior *et al.,* 2016). However, when the data set of independent variables presents a greater quantity of samples and/or has multicollinearity, the method is not successful and presents deviations (Das *et al.,* 2018).

Diseases and pests have been causing a reduction in coffee quality and yield in Brazil (Spongoski *et al.,* 2005). There are few studies on pest and disease indices in coffee trees, not only in Brazil, but in all countries where the plant is cultivated. The most common strategy to control these diseases and pests is the application of foliar fungicides and insecticides, depending on their intensity in the region. Unfortunately, this traditional method does not consider the influence of climatic conditions on the development of diseases and pests. With the use of these algorithms of machine learning, computers can make accurate decisions of the best moments of application of agrochemicals for the control of pests and diseases.

The hypothesis of this study is that one of the prediction algorithms calibrated for field conditions will be simple to understand, easy to use, and accurate, so that it can be used to develop a disease and pest warning system to improve the efficiency of chemical control of pests of coffee plants.

Thus, the objective of this study was: a) To evaluate progress curves of the incidence of pests and diseases for different regions and the load of fruit present on plants; b) Correlate infection rates of pest and disease incidence with meteorological variables considering different time periods for each condition; c) Calibrate different algorithms to predict the incidence of pests and diseases based on climate, using machine learning algorithms including the most conventional methods (RLM and kNN) and the most flexible ones (RFR and RNA); d) Test the algorithms with independent data to assess their suitability for use in a disease and pest alert system, and, e) spatialize the accuracy of the best prediction algorithm for the incidence of pests and diseases for the Minas Gerais state.

Material and methods

We used a historical series of climatic data and levels of pest infestation and disease severity of *Coffee arabica* L. for the state of Minas Gerais, Brazil. We use representative localities of the State for coffee production: Boa Esperança (BOE), Carmo de Minas (CDM), Muzambinho (MUZ) and Varginha (VAR), located in the southern region of Minas Gerais (SO_{MG}) and Araxá (ARX), Araguari (ARG) and Patrocínio (PAT) located in the Cerrado Mineiro region (CE_{MG}) (Table 1). The area selected for the work corresponds to 875,060 ha.

TABLE 1. Geographic characteristics of the main coffee growing areas of the State of Minas Gerais.

	Southern Minas Gerais				Cerrado Mineiro (Mineiro Triangle and Upper Paranaíba)		
Places	Boa Esperança	Carmo de Minas	Muzambinho	Varginha	Araguari	Araxá	Patrocínio
Latitude (º)	21º 03' 59" S	22º 10' 31" S	21º 20' 47" S	21º 34' 0" S	18º 59' 35" S	19º 33' 21" S	18º 33' 21" S
Longitude (º)	45º 34' 3" W	45º 09' 0" W	46º 32' 0" W	45º24'2" W	46º 59' 0" W	46º 58' 0" W	48º 12' 25" W
Altitude (m)	830	1080	1033	940	961	960	933
Period	2010-2018	2006-2018	2010-2018	1998-2018	2010-2018	2010-2018	2010-2018
Area (km^2)	860,7	323,3	409	395,6	2 731	1 165	2 866

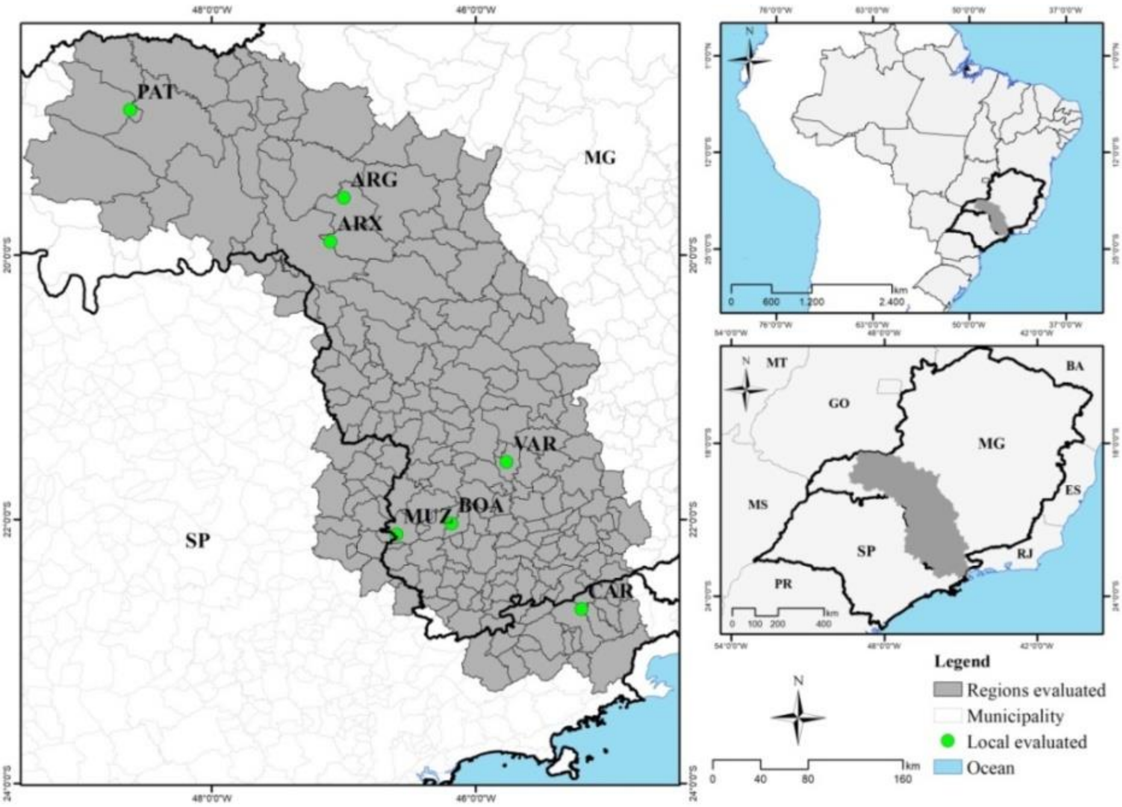

FIGURE 1. Coffee regions utilized in the study in the state of Minas Gerais, Brazil. BOE is Boa Esperança, CAR is Carmo de Minas, MUZ is Muzambinho, VAR is Varginha, ARG is Araguari, ARX is Araxá and, PAT is Patrocínio.

The meteorological data were obtained from automatic meteorological stations on a daily scale: minimum air temperature (T_{AIR}, ºC), maximum air temperature (T_{AIR}, ºC), precipitation (P, mm) and relative humidity (%).The weather stations were of type Vantage Pro2 Davis (K6162) and were installed near the coffee plantations that were evaluated.

With the meteorological elements measured by the automatic meteorological stations, other specific variables that influence the diseases and pests of the coffee tree were also selected (Table 2). The meteorological data were organized in matrix form (types of variables x time) to be compatible with the rate of infection of diseases and pests.

TABLE 2. Weather variables used in evaluations levels of pest infestation and disease severity of Coffee arabica.

Climatic acronyms	Definition
Tmin	Minimum temperature average (°C)
Tmax	Maximum temperature average (°C)
Rainfall	Total rainfall (mm)
NDR≥1mm	Number of days with rainfall ≥ 1 mm and < 9 mm
NDR≥10mm	Number of days with rainfall ≥ 10 mm
RH	Average relative humidity (%)
$NdRH_{90\%}$	Number of days with relative humidity ≥ 90%
$NdRH_{80\%}$	Number of days with relative humidity ≥ 80%

The analyses were performed considering three different time periods: 1-10d (from 1 to 10 days before the incidence evaluation); 11-20d (from 11 to 20 days before the incidence assessment); and 21-30d (from 21 to 30 days before the incidence assessment). These periods were selected by analyzing the latency period (time between infection of the pathogen in the plant and the manifestation of disease symptoms). For example, the latency period of Cercosporiose is 9 to 15 days and that for rust varies from 25 to 40 days (Kushalappa *et al.,* 1983).

Data on disease severity and pest infestation were provided by the Procafé Foundation (http://fundacaoprocafe.com.br/) from evaluations in fields with no phytosanitary treatment in the localities of Table 1. The disease data were: coffee rust (*Hemileia vastatrix*) and cercospora (*Cercospora coffeicola*) and pest data: coffee miner (*Leucoptera coffeella*) and coffee borer (*Hypothenemus hampei*). These data were collected from high and low slope crops.

Incidences were measured in a non-destructive procedure. The plants were randomly chosen in a zigzag walking pattern in the area, as recommended by Chalfoun (1997). Incidence scores were determined on leaves from the third or fourth knot of branches at the middle third of the plants. The site of collection of the pest/disease in the plant was according to Figure 2 and the methodology for evaluating coffee development, diseases, and pests used was according to Table 3.

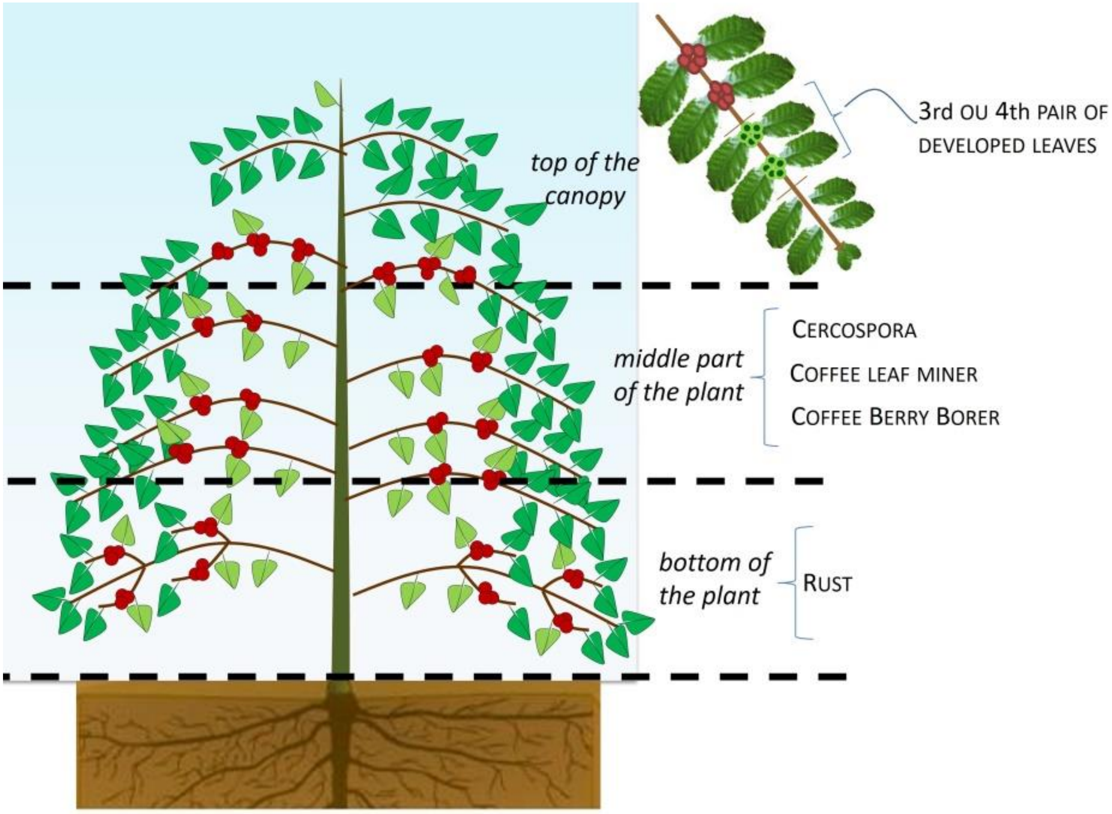

FIGURE 2. Place of greatest occurrence of pests and diseases in a coffee tree and lateral branches of coffee tree, with indication of the 3rd or 4th pair of leaves.

TABLE 3. Methodology for evaluating coffee development, diseases, and pests used by the Procafé Foundation of Brazil.

Phytosanitary problem	Methodology
Phenology	
Number of nodes	- Sample 20 plants per plot (random) - Select four branches per plant in the middle third (one on each side) - Quantify the number of nodes developed from September in each chosen branch
Leafiness of coffee plant	- Sample 20 plants per plot (random) - Select four branches per plant in the middle third (one on each side) - Quantify the amount of foliar development in the chosen branches - Quantify the soil-foliage percentages in the samples
Diseases	
Coffee rust (*Hemileia vastatrix*) and Cercospora index (*Cercospora coffeicola*)	- Sample 20 plants per field - Collect the leaves in the middle third of the chosen plant - Choose five lateral branches at random on each side of the plant - Remove a completely developed leaf, of the 3rd or 4th pair of leaves, from each branch - Total of 10 leaves/plant (five on each side) - Total of 200 leaves/field - Quantify the percentage of disease in the samples
Pests	
Coffee leaf miner (*Leucoptera coffeella*)	- Sample 20 plants per field - Collect the leaves in the middle third of the chosen plant - Choose five lateral branches at random on each side of the plant - Remove a completely developed leaf, of the 3rd or 4th pair of leaves, from each branch - Total of 10 leaves/plant (five on each side) - Total of 200 leaves/field - Quantify the percentage of the pests in the samples
Coffee berry borer (*Hypothenemus hampei)*	- Sample average of 50 plants per plot - Choose four branches per plant (one on each side) - Collect 25 fruits/branch for a total of 100 fruits/plant - 50 plants/field for a total of 5000 fruits/field - Quantify the percentage of the pest in the samples

To quantify coffee leafage (%), equation 2 was used. It is normal for each node developed from the plagiotropic branch of the coffee tree to present 2 leaves (1 leaf on each side of the node). To quantify the % of diseases and pests, equations 3-6 were used.

$$Leafiness\ (\%)\ =\ [\frac{(\frac{LEAF\ NUMBER}{2})}{Number\ of\ nodes}]\ \cdot\ 100 \qquad (2)$$

$$CLM\ (\%)\ =\ [\frac{(Number\ of\ leaves\ with\ live\ larvae)}{Number\ of\ leaves\ total}]\ \cdot\ 100 \qquad (3)$$

$$CBB\ (\%)\ =\ [\frac{(Number\ of\ fruits\ with\ berry\ borer)}{Number\ of\ fruits\ total}]\ \cdot\ 100 \qquad (4)$$

$$FI\ (\%)\ =\ [\frac{(Number\ of\ leaves\ with\ Ferrugem)}{Number\ of\ leaves\ total}]\ \cdot\ 100 \qquad (5)$$

$$CI\ (\%)\ =\ [\frac{(Number\ of\ leaves\ with\ Cercospora)}{Number\ of\ leaves\ total}]\ \cdot\ 100 \qquad (6)$$

where, Ci is Cercospora index (%), FI is Coffee rust index (%), CBB is Coffee berry borer (%) and CLM is Coffee leaf miner (%).

The diseases and pests of the coffee plantations were evaluated in two situations of "high" and "low" productivity, which occur due to the natural biennial nature of the coffee plant. “High” represents more than 30 bags of 60 kg ha^{-1}, while “low” is lower than 10 bags of 60 kg ha^{-1}. The disconnected range between these classifications is due to the difference observed in the field during subsequent high and low production seasons. Catuaí and Mundo Novo were the plants used, and both are susceptible to diseases and pests.

Pest infestation and severity of coffee diseases were evaluated seasonally from vegetative growth (September) and were represented by boxplot graphs, which allow for the description of the evolution and variability of pest diseases over time (Nutter, 1997; Bergamin Filho, 2011).

We sought to detail the relationship between coffee diseases and pests and the meteorological elements using univariate correlations (Pearson) of each element with the diseases for the south of Minas Gerais and Cerrado in both pending loads. Correlations were made considering three different time periods: 1-10d (from 1 to 10 days before the incidence evaluation); 11-20d (from 11 to 20 days before the incidence evaluation); and 21-30d (from 21 to 30 days before the incidence evaluation).

We used different methodologies to predict coffee disease and pest indices. The infection rate of diseases and pests was the dependent variable and the meteorological elements (Table 2) the independent variables of the models. In all methodologies, 40% of the data for the training and 60% for the calibration of the models were separated using python's library (sklearn.model_selection.train_test_split).

The forecasting methods were as follows: 1) Multiple Linear Regression (RLM); 2) K Neighbors Regressor (KNN); 3) Random Forest Regressor (RFT) and Artificial Neural Networks - Multi-layer Perceptron (MLP).

We use the ridge method in RLM. This method avoids poor conditioning of the matrix of the regressors variables, controlling the inflation and the general instability found in least squares estimators. Briefly, it avoids the multicollinearity problem without having to exclude regressor variables, so it has no information loss. The KNN algorithm is a simple and easily implemented technique and is very flexible. In the KNN the 3 closest neighbors were identified, and the metric used to calculate the distances was the Euclidean distance. In the RFT, a forest was randomly created with an ensemble of 100 decision trees to predict pests and diseases as a function of the climate. The artificial neural network used was the Multi-layer Perceptron (MLP), with 3 layers of neurons. In each of these layers 10 neurons were used (hidden_layer_sizes = 10, 10, 10). MLP training was done using back propagation.

The actual observed field data and the results of all models were compared using several statistical indices: accuracy, precision, and level of significance (Table 4). Accuracy indicates the closeness of an estimate to the observed value and was evaluated using the Willmott's 'd' and the root mean square error (RMSE). Precision is the ability of a model to repeat an estimate and was evaluated using the coefficient of determination (R^2) adjusted (R^2adj) following Cornell and Berger (1987).

TABLE 4. Accuracy and precision of statistical indices that were used in the evaluation of models forecasting. YEST is estimated value of y; YOBS is observed value of y; XOBS is observed value of x; N is number of data.

Statistical Index	**Equation**
Accuracy	
Willmott's 'd'	$d = 1 - \frac{\sum_{i=1}^{N}(Yobs_i - Yest_i)^2}{\sum_{i=1}^{N}\left(\left\|Yest_i - \bar{Y}\right\| + \left\|Yobs_i - \bar{Y}\right\|\right)^2}$
RMSE	$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(Y_{OBS_i} - Y_{EST_i})^2}{N}}$
Precision	
R^2	$R^2 = 1 - \frac{\sum_{i=1}^{n}(Yest_i - \bar{Y}obs)^2}{\sum_{i=1}^{n}(Yest_i - \bar{Y}obs)^2 - \sum_{i=1}^{n}(Yest_i - Yobs_i)^2}$

With all the algorithms calibrated for the prediction of pests and diseases, the RMSE index of the most accurate algorithm for the south of Minas and Cerrado Mineiro regions was mapped in both pending production loads. Maps were generated using kriging interpolation (Krige, 1951) with 1 neighbour and a resolution of 0.25° in the spherical model.

Results and discussion

The southern region of Minas Gerais (SO_{MG}) showed an average annual air temperature (T_{AIR}) 1.39ºC lower than the values measured in the Cerrado Mineiro (CE_{MG}). The rainfall (P) was more similar, since the annual mean values were 116.5 and 120.5 mm mo^{-1} for SO_{MG} and CE_{MG}, respectively (Figure 3.A.B). The seasonal distributions of T_{AIR} and P in the SO_{MG} and CE_{MG} regions were similar. In May to August, the lowest T_{AIR} and P occurred in both regions, with mean values of 17.08 ºC and 30.01 mm in the SO_{MG} and 18.8 ºC and 23.9 ºC in the CE_{MG}. Results were similar to those observed by Aparecido *et al.,* (2018).

The lowest water storage in the soil occurs in August in the SO_{MG} and in September in the CE_{MG} (Figure 3.G.H). The water deficit (DEF) occurs in both regions from April to November, however, the DEF is more intense in the CE_{MG} region, reaching -35.8 mm in August (Figure 3.I.J). The annual cumulative DEF was -65.7 and -103.6 mm y^{-1} for the SO_{MG} and CE_{MG}, respectively. DEF is one of the variables that most influences the development of agricultural crops, as well as that of coffee (DaMatta, 2004; Carvalho *et al.,* 2011; Syvertsen and Garcia-Sanchez, 2014).

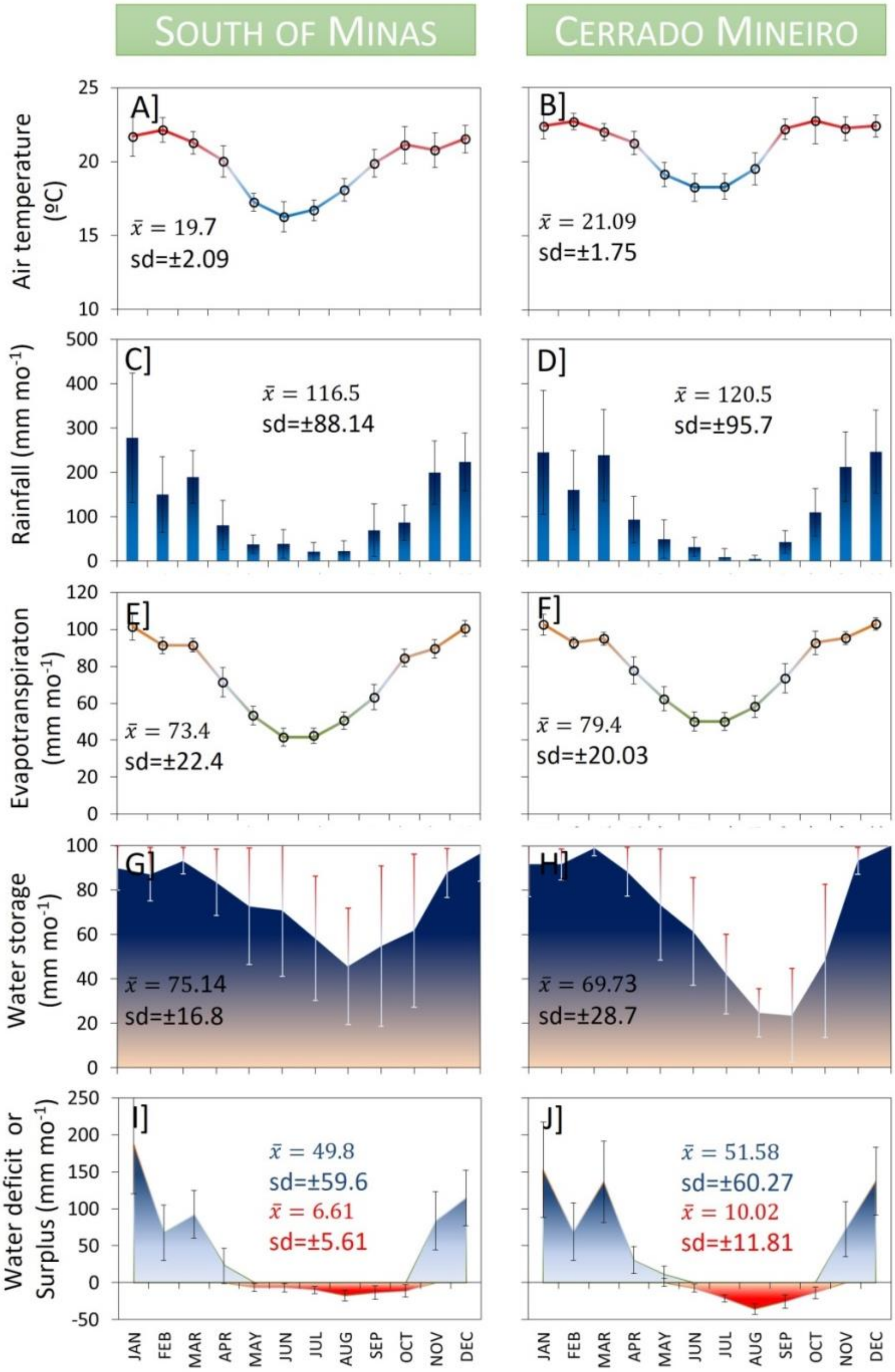

FIGURE 3. Air temperature (ºC), rainfall (mm), evapotranspiration (mm), water storage (mm) and water deficit or water surplus (mm) of the regions in South of Minas Gerais and Cerrado Mineiro from the period 1995 to 2018, Minas Gerais, Brasil. Legend: $\bar{\bar{x}}$ is the annual mean of the variable and *sd* is the annual standard deviation of the climatic variable.

Coffee rust showed a trend of sigmoidal progression in both slopes and in all studied regions (Figure 4 and 5). SO_{MG} showed a mean overall incidence of coffee rust of 23.77% (± 24.13) and CE_{MG} an average incidence of 24.46% (± 28.40) of infected leaves. The lowest values of rust occurred between September and December (4%). After April there was a stabilization of growth in SO_{MG} and CE_{MG}. This tendency of growth of coffee rust curves was also shown by Hinnah (2019). In the SO_{MG}, the coffee with high yield had higher rust indices (31.46%, ± 30.92) in relation to coffee with low yield (16.08%, ± 17.35). In the period from May to July, the coffee in the SO_{MG} with high yield had 73.5% rust incidence, while the low-yield coffee had average values of 24.8% of infected leaves (Figure 4). In the CE_{MG} there were no differences between the high and low coffee loads. For high yield the values of rust were 24.87% (± 29.55) and for low yield the values were 24.05% (± 27.25) of infected leaves.

The locality with the lowest incidence of coffee rust in the SO_{MG} was Boa Esperança, with annual mean values of 28.05% (± 30.47) and 10.74% (± 12.52) for high and low yields, respectively. Varginha was the locality with the highest incidence, with 34.41% (± 33.24) for high yield and 17.76% (± 19.11) for low yield. Patrocínio was the locality of CE_{MG} with the lowest average incidence of coffee rust at high yield (19.72%, ± 25.57) and with the highest incidence at low yield (24.63%, ± 25.61). Araxá had the highest incidence of rust at high yield (28.20%, ± 31.93) and lower incidence at low yield (23.01%, ± 28.60) of infected leaves.

The control of rust in coffee can be made with a protective fungicide (copper base) when the incidence levels reach a maximum of 5%. If the incidence level exceeds 5% and reaches up to 12%, it is necessary to apply systemic fungicides. Thus, it is necessary to engage in control of the disease in both regions and crop loads until November using protective fungicides and from December using systemic fungicides, so that the diseases do not reach uncontrollable levels. Rust, if not controlled correctly, causes early leaf fall and subsequent drying of the productive branches of the crop, which negatively reflects on the development of flower buds and fruit growth, and causes a drastic reduction of the productivity of the next agricultural year. Another way to control rust is the use of resistant cultivars such as 'Obatã IAC 4739', 'Catucaí Amarelo 24/137' or 'Obatã IAC 1669-20' (Carvalho et al., 2012; Avelino et al., 2015; Fazuoli et al., 2018).

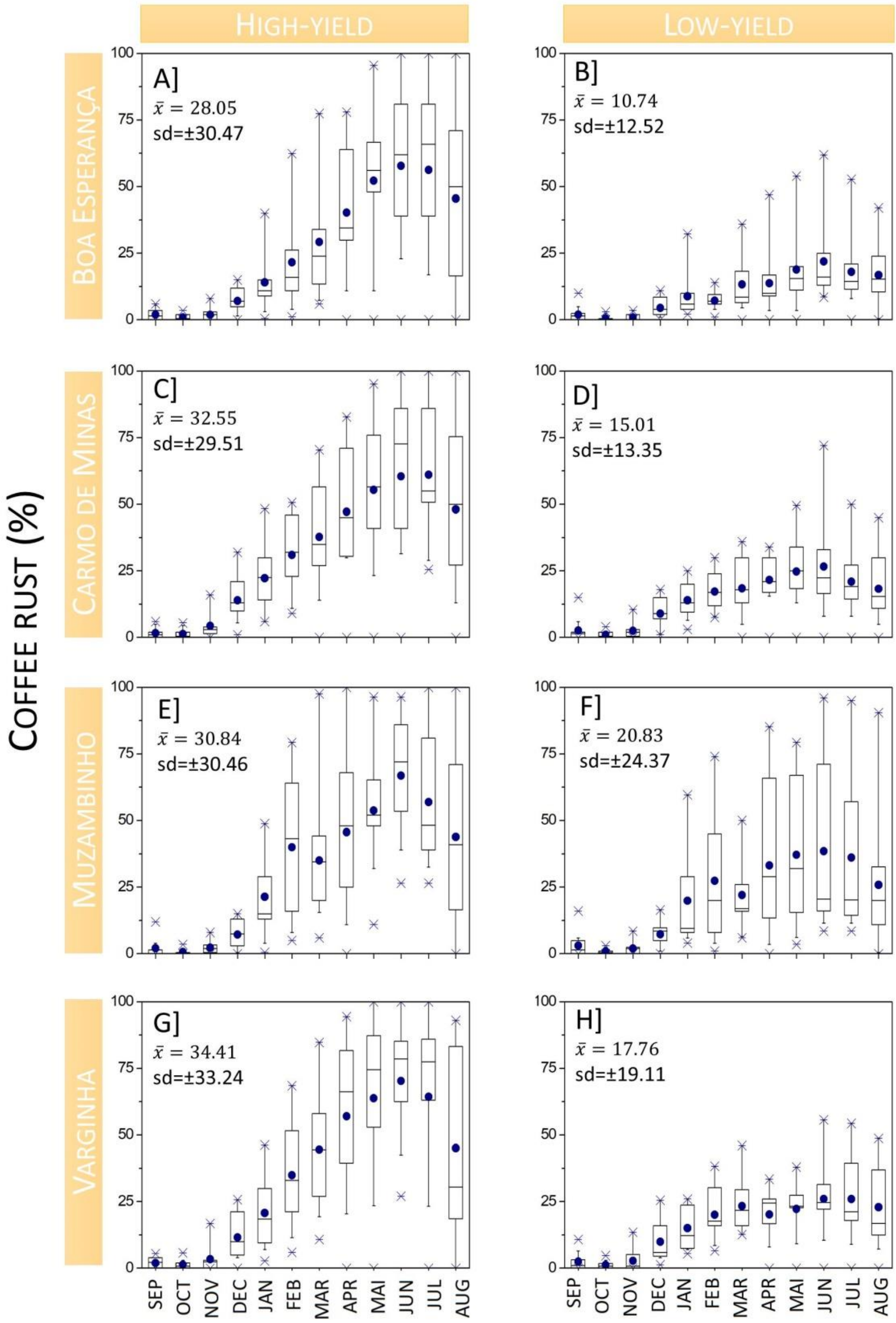

FIGURE 4. Coffee-rust incidence progress in two different field conditions (high and low yield) in Boa Esperança (A-B); Carmo de Minas (C-D); Muzambinho (E-F) and Varginha (G-H).

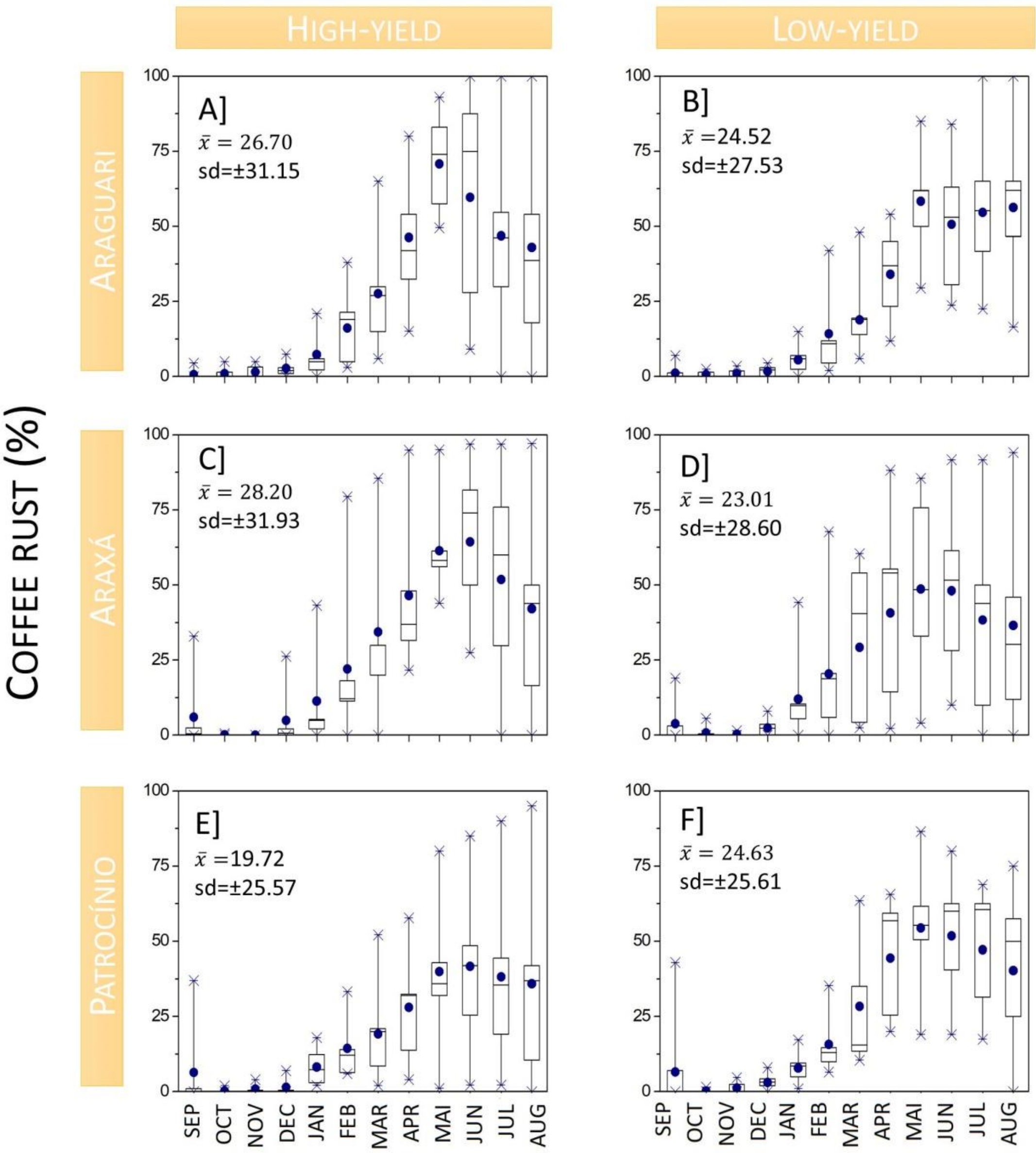

FIGURE 5. Coffee-rust incidence progress in two different field conditions (high and low yield) in Araguari (A-B); Araxá (C-D) and Patrocínio (E-F).

The cercospora of the coffee had a tendency of sigmoidal progress in both regions and crop loads (Figure 6 and 7). SO_{MG} showed a mean overall incidence of cercospora of only 4.26% (± 4.63), while in CE_{MG} the mean incidence was higher, with values of 17.32% (± 18.68) of infected leaves. The highest incidences started from December with stabilization after February, for both SO_{MG} and CE_{MG}.

In the SO_{MG}, the high-yield coffee had cercospora rates similar to the low-yield, with values of 4.68% (± 5.09) and 3.84% (± 4.29) of infected leaves, respectively. At CE_{MG} there were no differences between production loads. The high-yield coffee had an incidence of cercospora of 18.22% (± 20.18) and in low-yield it was 16.41% (± 17.1) of infected leaves.

The locality with the lowest incidence of cercospora of the coffee tree in the SO_{MG} was Varginha, with annual mean values of 3.74% (± 4.03) and 2.18% (± 2.47) for high and low-yield coffee, respectively. And the locality with the highest average incidence was Muzambinho, with 5.84% (± 6.0) for high-yield and 6.02% (± 6.72) for low-yield.

Patrocínio was the locality of CE_{MG} with the lowest mean incidences of cercospora in high and low-yields, with values of 8.25% (± 9.82) and 5.15% (± 5.01), respectively (Figure 7.E.F), and Araguari presented the highest incidence of cercospora in high and low-yields, with values of 32.91% (± 32.74) and 30.90% (± 31.38), respectively.

The control of the cercospora of the coffee should be made with protective fungicides, mainly copper based, as well as of systemic fungicides, but both have a greater effectiveness if they are applied as preventative measures, and in this case the forecast models can facilitate decision-making by farmers.

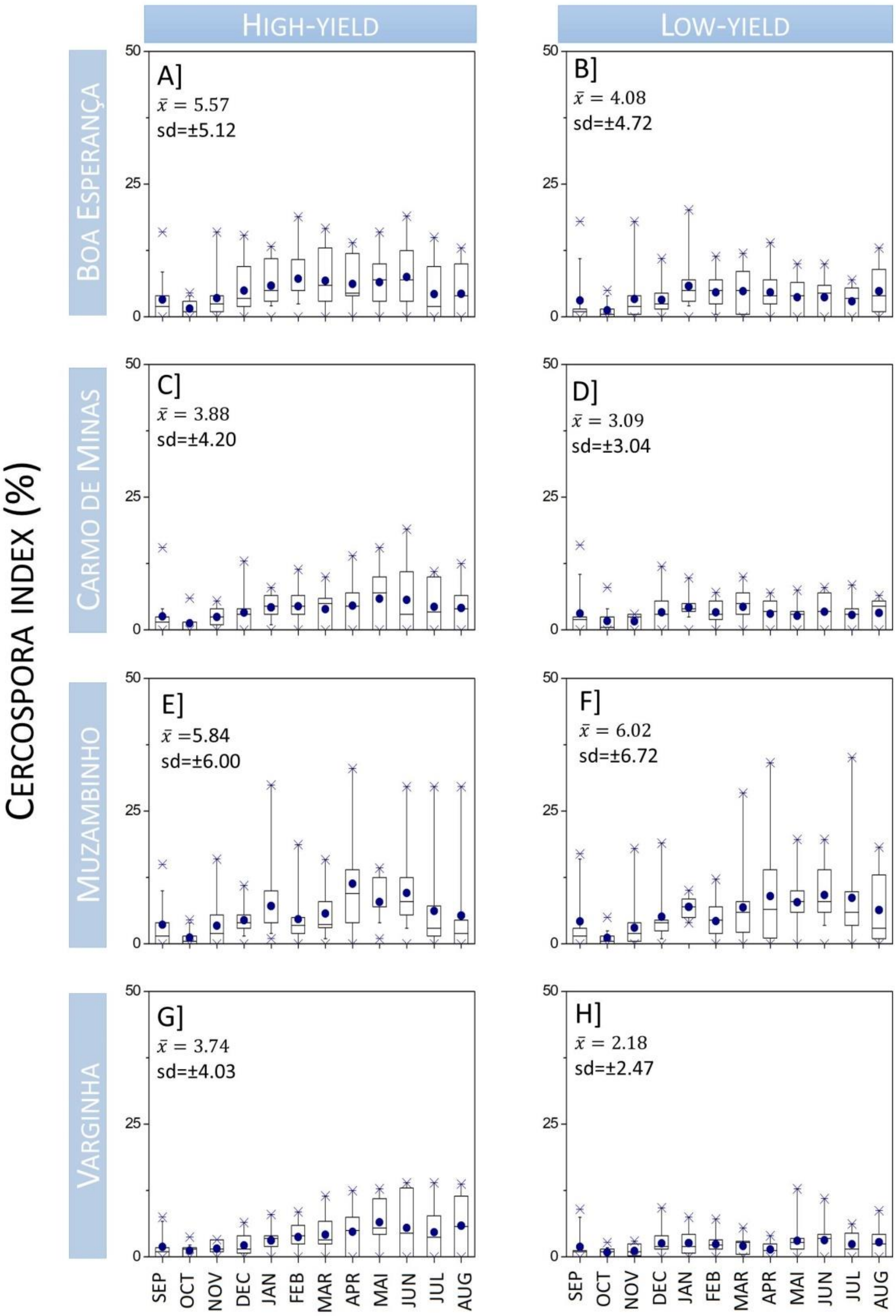

FIGURE 6. Cercospora incidence progress in two different field conditions (high and low yields) in Boa Esperança (A-B); Carmo de Minas (C-D); Muzambinho (E-F), and Varginha (G-H).

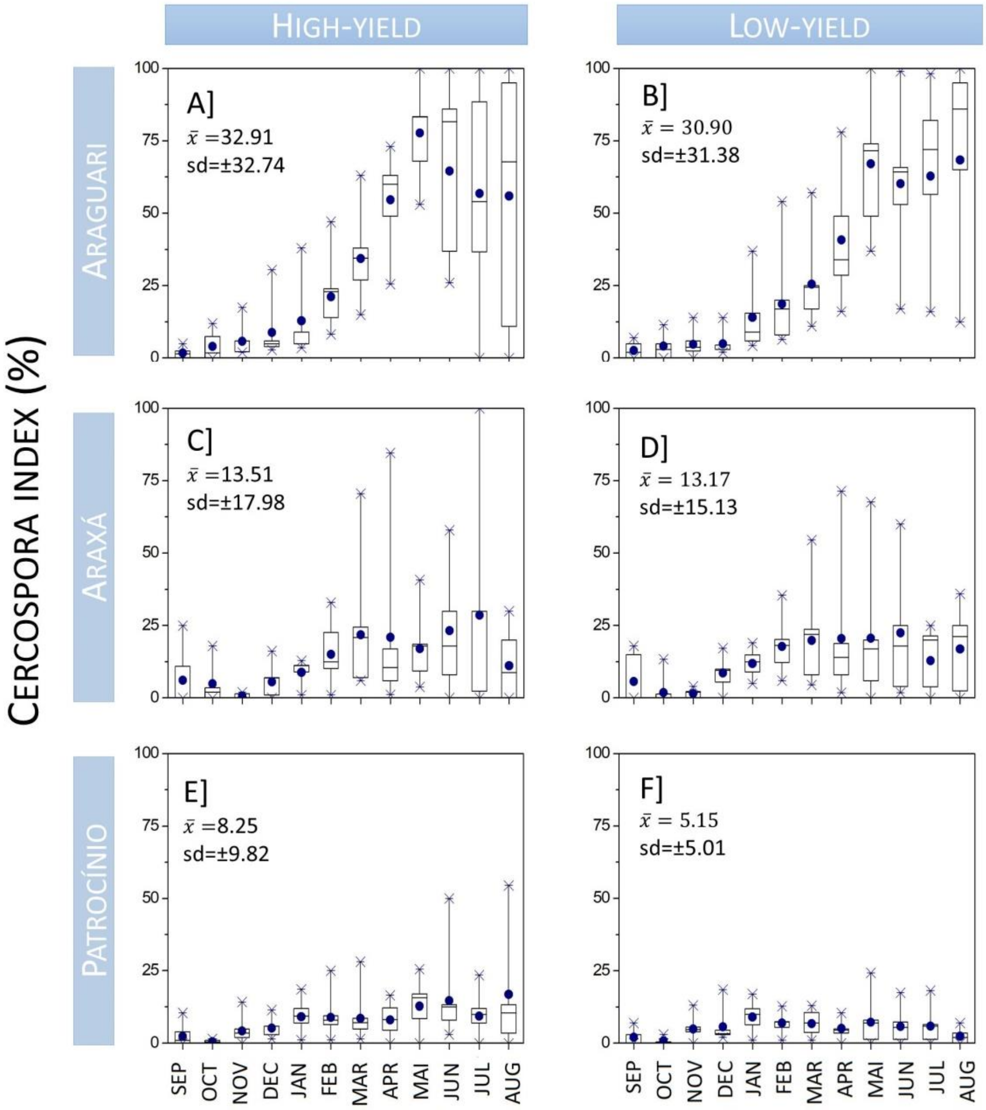

FIGURE 7. Cercospora incidence progress in two different field conditions (high and low yield) in Araguari (A-B); Araxá (C-D) and Patrocínio (E-F).

The coffee miner showed progress curves with a sigmoidal tendency for both regions and crop loads (Figure 8 and 9). The SO_{MG} showed a general average incidence of coffee miner of only 1.55% (± 3.34), while in the CE_{MG} the average incidence was higher, with values of 12.83% (± 16.83) of infected leaves.

The lowest incidences of the coffee miner were from October to January in both loads and for all regions, and this occurs due to the onset of rainfall. This result was also highlighted by Conceição *et al.* (2005) who observed a reduction in the population of the miner in the period with high rainfall rates.

The high-yield coffee in the SO_{MG} had rates of coffee miner similar to the coffee with low yields, with values of 1.55% (± 3.98) and 1.56% (± 2.70) of infected leaves, respectively. In the CE_{MG} the pest severity was higher in both coffee loads, with values of 12.66% (± 16.72) and 13.00% (± 16.94) of infected leaves for the high and low-yields, respectively.

The locality with the lowest incidence of Coffee miner in SO_{MG} was Varginha, with annual average values of 0.84% (± 1.45) and 1.06% (± 1.61) for high and low-yields, respectively (Figure 8.G.H), and the locality with the highest average incidence was Muzambinho, with 2.61% (± 10.25) for high and 2.07% (± 4.03) for low-yields.

Araxá was the locality of the CE_{MG} with the lowest average incidence of coffee miner in high and low yields, with values of 3.43% (± 7.43) and 4.88% (± 10.42), respectively (Figure 9.C.D). The location with the highest indexes was Araguari with values of 20.55% (± 27.85) and 20.69% (± 28.01), respectively (Figure 9.A.B). These higher levels of coffee miner occur due to higher air temperatures in CE_{MG}, thus shortening the miner's cycle. Thus, the shorter the duration of the cycle, more generations occur in a short time, resulting in extremely high populations of larvae and adults.

The control of the coffee miner can be done with insecticides and should be initiated when the infestation reaches the control level (20 to 30%), or more, of mined leaves in the upper third of the coffee trees. Thus, it is necessary to engage in pest control in all CE_{MG} locations, especially during the period of December-January (Figure 9). In SO_{MG} the control of the pest is shown to be sporadic and varies by location. Control is necessary because the Coffee miner promotes drastic defoliation that influences flowering and fruit formation (Guerreiro-Filho, 2006).

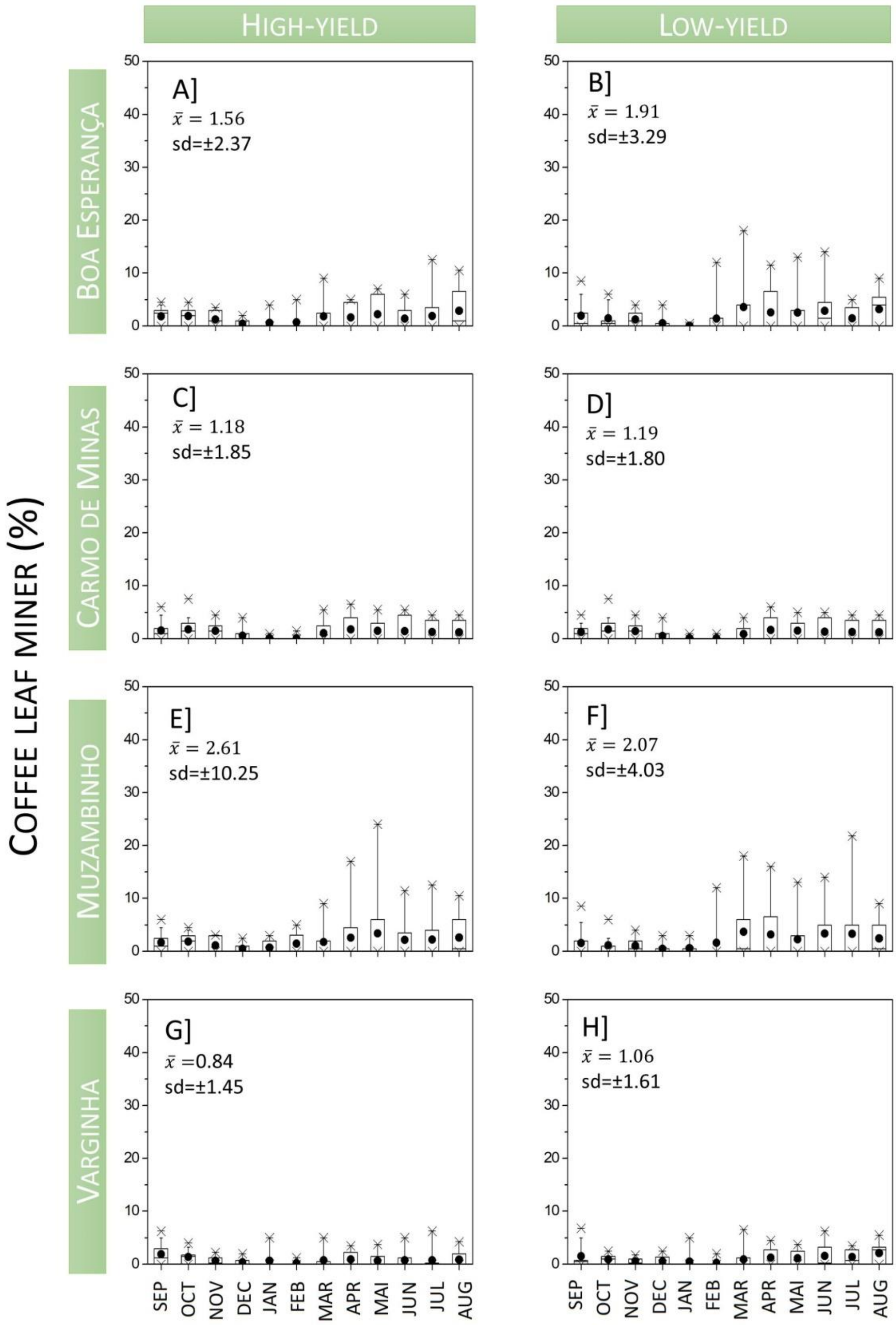

FIGURE 8. Coffee leaf miner incidence progress in two different field conditions (high and low yield) in Boa Esperança (A-B); Carmo de Minas (C-D); Muzambinho (E-F) and Varginha (G-H).

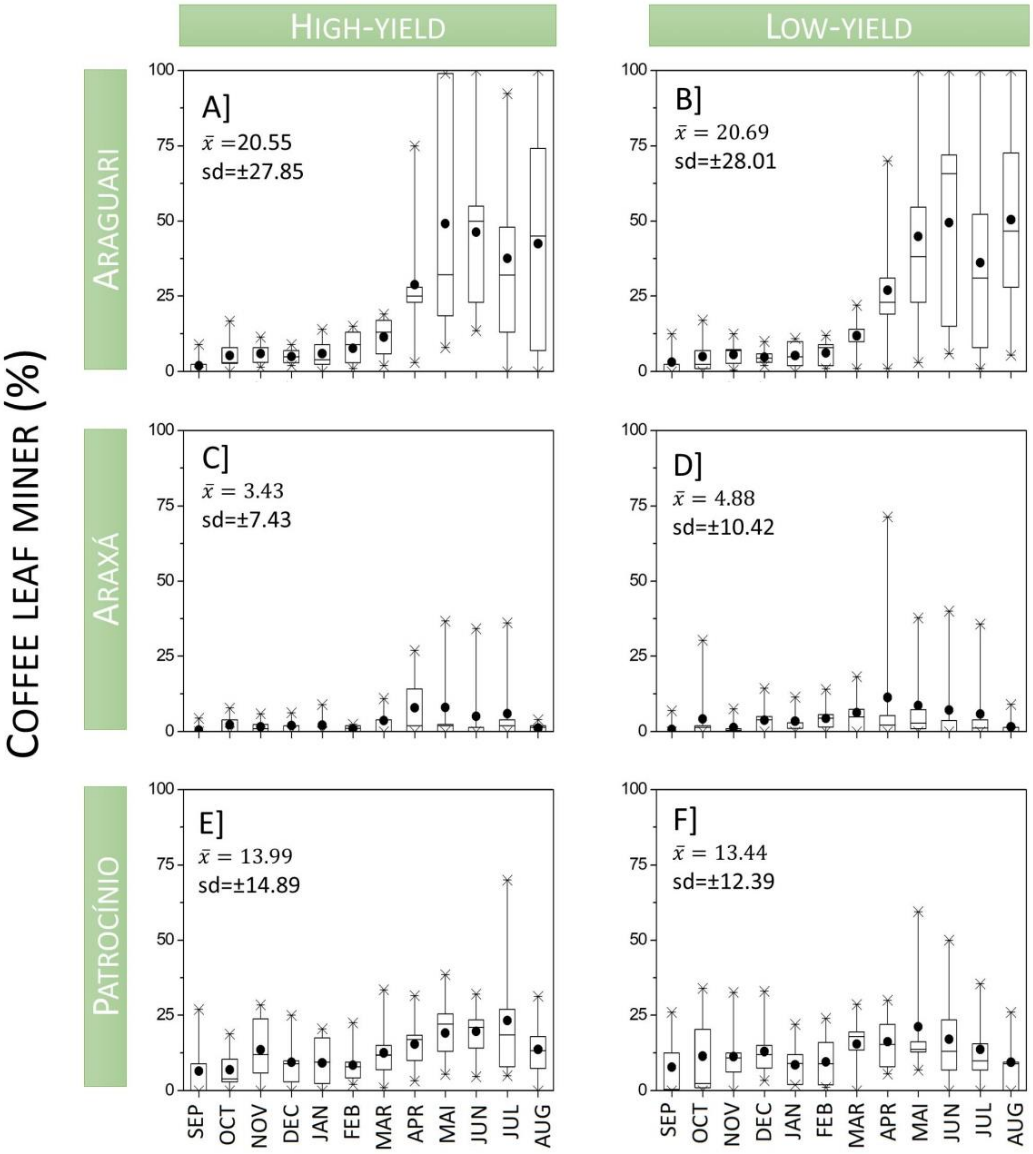

FIGURE 9. Coffee leaf miner incidence progress in two different field conditions (high and low yield) in Araguari (A-B), Araxá (C-D) and Patrocínio (E-F).

The coffee borer showed Gaussian trend curves in both regions and productive loads (Figure 10 and 11). SO_{MG} showed a mean average incidence of the coffee borer of only 0.50% (± 1.33), while in CE_{MG} the mean incidence was higher, with values of 2.57% (± 6.03) of infected leaves. The peak incidence of the pest occurred between December and May, for both regions.

In the SO_{MG} the high-yield coffee showed similar coffee borer indices to those for the low-yield coffee, with values of 0.57% (± 1.49) and 0.43% (± 1.16) of infected leaves, respectively. In CE_{MG} the coffee borer rates were higher in both coffee loads, with 2.34% (± 5.39) and 2.80% (± 6.65) of infected leaves at high and low-yields, respectively.

The locality with the lowest incidence of coffee borer in the SO_{MG} was Varginha, with annual average values of 0.38% (± 0.80) and 0.18% (± 0.44) for high and low-yields, respectively (Figure 10.G.H), and the locality with the highest average incidence was Carmo de Minas, with 0.77% (± 2.18) for high-yields and 0.56% (± 1.52) for low-yields.

Araxá was the locality of the CE_{MG} with the lowest average incidence of coffee borer in high and low-yields, with values of 0.67% (± 2.71) and 2.19% (± 6.91), respectively (Figure 11.C.D). Araguari had the highest incidence of coffee borer in high and low-yields, with values of 4.86% (± 10.13) and 4.52% (± 9.50), respectively.

The control of the coffee borer should be done with insecticides and should be started when the infestation reaches the control level (3% to 5%), spraying the most attacked parts of the crop (Souza et al., 2013b). Thus, it is necessary to engage in control of the pest in all the localities of CE_{MG}, mainly in the period of January-April (Figure 11). In the SO_{MG} the control of the coffee borer is shown to be sporadic and varies by location.

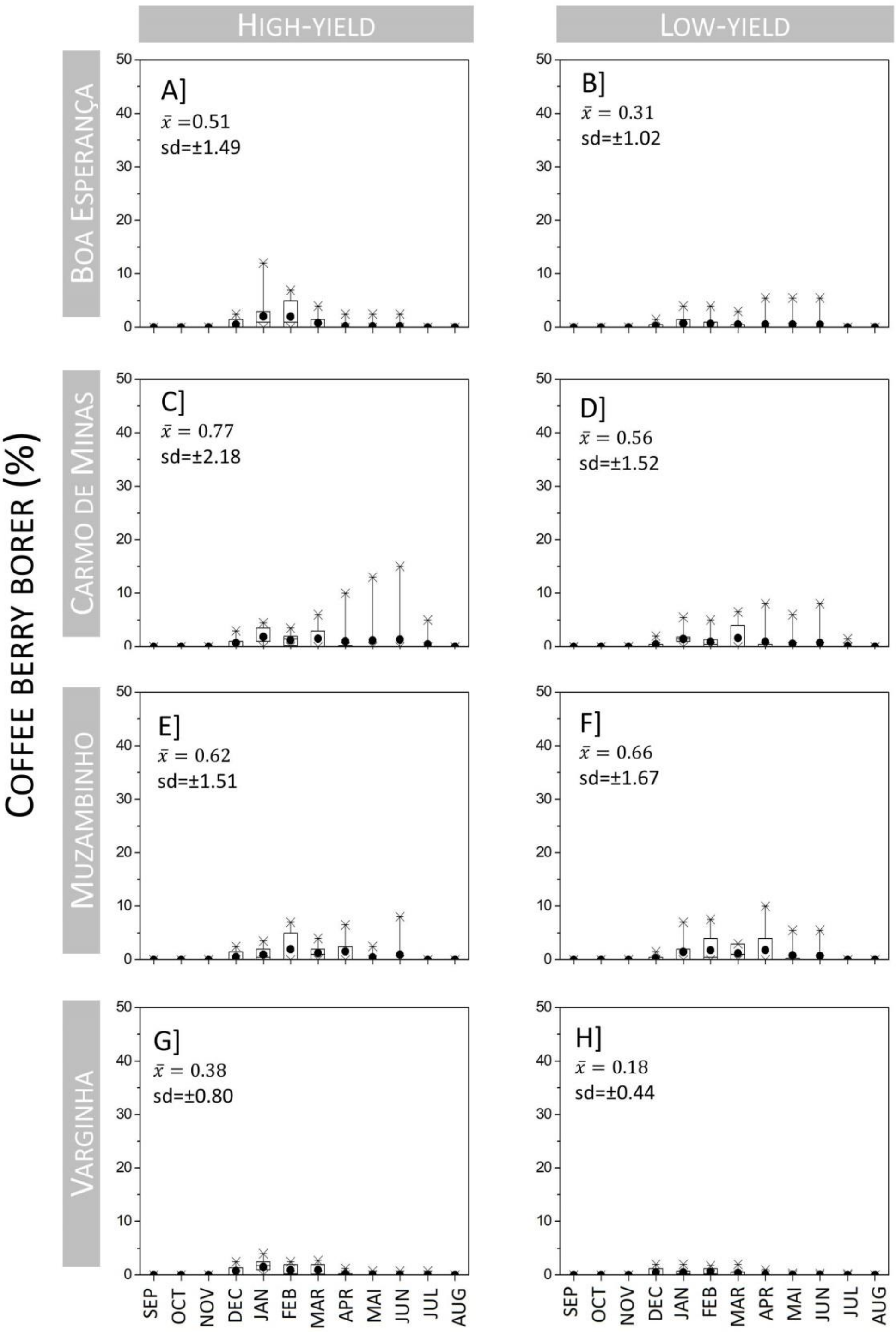

FIGURE 10. Coffee berry borer incidence progress in two different field conditions (high and low yield) in Boa Esperança (A-B); Carmo de Minas (C-D); Muzambinho (E-F) and Varginha (G-H).

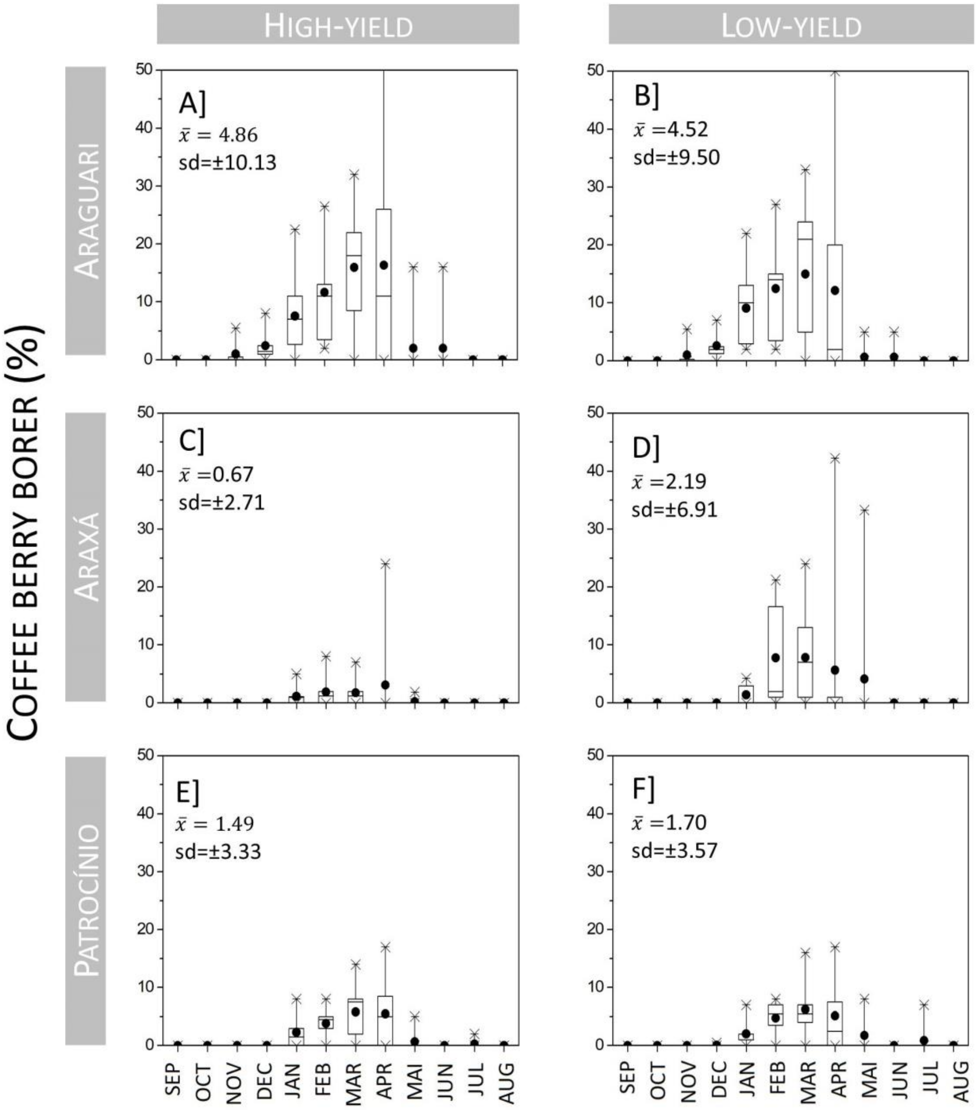

FIGURE 11. Coffee berry borer incidence progress in two different field conditions (high and low yield) in Araguari (A-B); Araxá (C-D) and Patrocínio (E-F).

The maximum air temperature (T_{MAX}), the number of days with relative humidity above 80% (NdRH> 80%) and relative humidity (RH) were the climatic variables that showed significant correlations with coffee rust disease (Table 4). T_{MAX}, NdRH> 80% and HR showed positive correlations and in all the periods 1-10d, 11-20d and 21-30d in SO_{MG}. For example, for the high-yield coffee in the south region the correlations were 0.541*, 0.48* and, 0.437* for periods 1-10d, 11-20d and 21-30d, respectively. These correlations demonstrate that an increase in the severity of rust occurs when the levels of T_{MAX}, NdRH> 80% and RH increase. Campbell and Madden (1990) and, Salgado *et al.,* (2007) report that T_{AIR} and RH are the variables that most influence the biological components of a pathosystem in the development of an epidemic.

TABLE 4. Pearson's correlation between monthly coffee rust infection rate and weather variables. Number of days before assessment of incidence of leaf symptoms: 1-10d is from 1 to 10 days; 11-20d is from 11 to 20 days; 21-30d is from 21 to 30 days. Minas Gerais, Brazil.

Weather variables	*Period assessed before symptoms*					
	1-10d	11-20d	21-30d	1-10d	11-20d	21-30d
	Southern Minas -High			Southern Minas - Low		
Tmax	0.541*	0.48*	0.437*	0.403*	0.326*	0.246*
Tmin	-0.197	-0.139	-0.005	-0.094	-0.025	0.09
NDR>1mm	-0.18	-0.012	0.103	-0.044	0.044	0.156
NDR>10mm	-0.154	-0.023	0.061	-0.082	0.013	0.065
NdRH>80%	0.333*	0.377*	0.393*	0.233*	0.257*	0.265*
NdRH>90%	0.2	0.272	0.304*	0.122	0.195	0.212
Rainfall	-0.162	-0.061	0.033	-0.044	-0.02	0.065
RH (%)	0.396*	0.437*	0.484*	0.325*	0.352*	0.357*
	Cerrado - High			Cerrado - Low		
Tmax	-0.491*	-0.431*	-0.439*	-0.478*	-0.409*	-0.406*
Tmin	-0.361*	-0.23	-0.07	-0.389*	-0.277	-0.102
NDR>1mm	-0.162	-0.046	0.058	-0.212*	-0.073	0.012
NDR>10mm	-0.173	-0.075	-0.018	-0.205	-0.098	-0.08
NdRH>80%	0.185	0.253*	0.298*	0.115	0.157*	0.244*
NdRH>90%	0.124	0.134	0.132	0.074	0.052	0.034
Rainfall	-0.223	-0.11	-0.006	-0.267	-0.143	-0.048
RH (%)	0.259	0.337*	0.383*	0.214*	0.291*	0.341*

(*) Asterisk indicates that the correlations are significant at p≤0.05.

The maximum air temperature (T_{MAX}) showed positive and significant correlations (p <0.05) with the incidence of cercospora in both productive loads and in both regions (Table 5). For example, in the SO_{MG} for high-yield coffee, the values were r = 0.281*, 0.308* and 0.246*, for periods 1-10d, 11-20d and 21-30d, respectively. The correlation shows that the higher the T_{MAX}, the higher the incidence of cercospora in coffee cultivation. This result has also been shown by several authors such as Echandi (1959), Zambolim *et al.,* (1997), Salgado et al. (2007) and Souza *et al.* (2013a) who reported that cercospora requires an excess of insolation and higher temperatures for the germination of fungi spores, occurring at 30°C. As a result of the increase in the severity of the cercospora, early defoliation occurs in the coffee plants, mainly due to the production of ethylene in the leaves that were injured, which promotes a reduction in their production (Zambolim *et al.,* 1997; Salgado *et al.,* 2007).

TABLE 5. Pearson's correlation between monthly cercospora infection rate and weather variables. Number of days before assessment of incidence of leaf symptoms: 1-10d is from 1 to 10 days; 11-20d is from 11 to 20 days; 21-30d is from 21 to 30 days. Minas Gerais, Brazil.

Weather variables	*Period assessed before symptoms*					
	1-10d	11-20d	21-30d	1-10d	11-20d	21-30d
	Southern Minas - High			Southern Minas - Low		
Tmax	0.281*	0.308*	0.246*	0.119	0.208*	0.131
Tmin	-0.032	-0.046	0.051	0.005	-0.043	0.039
NDR>1mm	-0.005	0.104	0.152	0.032	0.108	0.119
NDR>10mm	-0.006	0.1	0.167	0.048	0.063	0.103
NdRH>80%	0.288*	0.291*	0.339*	0.128	0.147	0.183
NdRH>90%	0.157	0.148	0.26	0.091	0.098	0.134
Rainfall	-0.083	0.098	0.123	-0.05	0.096	0.086
RH (%)	0.266	0.319*	0.313*	0.13	0.211*	0.182
	Cerrado - High			Cerrado - Low		
Tmax	0.343*	0.286*	0.261	0.322*	0.199	0.212
Tmin	-0.183	-0.106	0.025	-0.099	-0.025	0.08
NDR>1mm	-0.051	0.023	0.113	0.032	0.087	0.165
NDR>10mm	-0.062	-0.03	0.024	0.029	0.016	0.08
NdRH>80%	0.118	0.134	0.199	0.12	0.112	0.183
NdRH>90%	0.042	0.035	0.047	0.043	0.011	0.031
Rainfall	-0.041	-0.042	0.026	0.056	0.017	0.065
RH (%)	0.221	0.262*	0.289*	0.249*	0.233	0.264*

(*) Asterisk indicates that the correlations are significant at $p \leq 0.05$.

Precipitation (P) showed negative and significant correlations ($p < 0.05$) with the incidence of coffee miner in the period 1-10d in both loads and in the two regions (Table 6), clearly demonstrating the decrease in the severity of the coffee miner with the increase of P. This demonstrates that the insect needs prolonged periods of drought to promote high infestation levels in coffee. This result is confirmed by Machado *et al.,* (2014) and Costa *et al.,* (2015).

The T_{MIN} had positive correlations with the incidence of coffee miner in the period 11-20d, however, only in the SO_{MG}. Thus, the higher the T_{MIN} values, the greater the severity of the coffee miner. This positive correlation between T_{AIR} and the severity of the coffee miner was also confirmed by other authors such as Fernandes et al. (2009). The CE_{MG} did not have this relationship, since it is already a region with higher T_{AIR}. Higher levels of coffee miner are undesirable, because the mined leaves fall before leaves that have not been attacked, which causes a reduction of the active photosynthetic area and consequent drop in coffee production (Caixeta *et al.,* 2004).

TABLE 6. Pearson's correlation between monthly coffee miner infection rate and weather variables. Number of days before assessment of incidence of leaf symptoms: 1-10d is from 1 to 10 days; 11-20d is from 11 to 20 days; 21-30d is from 21 to 30 days. Minas Gerais, Brazil.

Weather variables	*Period assessed before symptoms*					
	1-10d	11-20d	21-30d	1-10d	11-20d	21-30d
	Southern Minas - High			Southern Minas - Low		
Tmax	0.023	0.107	0.093	0.123	0.206	0.193
Tmin	0.16	0.228*	0.183	0.162	0.218*	0.172
NDR>1mm	-0.245*	-0.198*	-0.139	-0.241*	-0.146	-0.076
NDR>10mm	-0.183	-0.149	-0.059	-0.173	-0.102	0.012
NdRH>80%	-0.059	-0.025	0.033	0.049	0.073	0.169
NdRH>90%	-0.004	-0.019	0.056	0.094	0.061	0.181
Rainfall	-0.234*	-0.154	-0.148	-0.251*	-0.119	-0.104
RH (%)	-0.119	-0.082	-0.064	-0.024	0.036	0.054
	Cerrado - High			Cerrado - Low		
Tmax	-0.095	-0.019	-0.006	-0.09	-0.029	-0.021
Tmin	-0.126	-0.079	-0.043	-0.084	-0.057	-0.027
NDR>1mm	-0.107	-0.10	-0.075	-0.075	-0.093	-0.011
NDR>10mm	-0.076	-0.132*	-0.093	-0.072	-0.106*	-0.044
NdRH>80%	-0.084	-0.124	-0.026	-0.018	-0.088	0.015
NdRH>90%	-0.023	-0.055	-0.051	-0.024	-0.042	-0.022
Rainfall	-0.122*	-0.107	-0.114	-0.117*	-0.084	-0.083
RH (%)	0.022	0.007	0.03	0.038	0.026	0.037

(*) Asterisk indicates that the correlations are significant at p≤0.05.

The T_{MIN} of the period 11-20d had positive and significant correlations ($p < 0.05$) with coffee borer incidence in both loads and regions (Table 7). For example, in the SO_{MG} for high-yield coffee the value was r=0.296. This correlation occurs because with higher T_{MIN} there is a reduction in the insect cycle and a consequent increase in the population of the coffee borer (Laurentino and Costa, 2004). Other climatic variables also show significant correlations with the level of pest infestation in both regions, such as RH in the periods 1-10d and 11-20d and NdRH> 80% in the period of 1-10d (Table 7).

TABLE 7. Pearson's correlation between monthly coffee berry borer infection rate and weather variables. Number of days before assessment of incidence of leaf symptoms: 1-10d is from 1 to 10 days; 11-20d is from 11 to 20 days; 21-30d is from 21 to 30 days. Minas Gerais, Brazil.

Weather variables	*Period assessed before symptoms*					
	1-10d	11-20d	21-30d	1-10d	11-20d	21-30d
	Southern Minas - High			Southern Minas - Low		
Tmax	-0.033	-0.05	0.068	-0.027	-0.033	0.046
Tmin	0.249	0.298*	0.253	0.237	0.293*	0.246
NDR>1mm	0.336*	0.306*	0.265	0.316*	0.282*	0.252
NDR>10mm	0.321*	0.289	0.154	0.333*	0.267	0.2
NdRH>80%	0.316*	0.267	0.129	0.312*	0.248	0.142
NdRH>90%	0.232	0.15	0.119	0.257	0.195	0.201
Rainfall	0.301	0.335*	0.235	0.26	0.324*	0.207
RH (%)	0.26*	0.28*	0.196	0.254*	0.263*	0.194
	Cerrado - High			Cerrado - Low		
Tmax	-0.15	-0.213	-0.123	-0.1	-0.144	-0.096
Tmin	0.212	0.295*	0.243	0.135	0.314*	0.18
NDR>1mm	0.329*	0.313	0.289	0.202	0.213	0.195
NDR>10mm	0.224	0.332	0.205	0.175	0.246	0.091
NdRH>80%	0.463*	0.333	0.271	0.32*	0.247	0.169
NdRH>90%	0.202	0.197	0.155	0.054	0.126	0.033
Rainfall	0.134	0.316	0.296	0.092	0.199	0.198
RH (%)	0.334*	0.361*	0.303	0.228*	0.233*	0.222

(*) Asterisk indicates that the correlations are significant at $p \leq 0.05$.

In predicting the incidence of rust the RFT algorithm demonstrated the highest accuracy (RMSE) and the highest precision (R2adj) in both fruit loads and in all locations (Table 8).

The average RMSE for the prediction of coffee rust in high-yielding crops was 19.59, 24.15, 8.51 and 15.48% for KNN, MLP, RFT and RLM, respectively. For low-yield coffee the RMSE values were 14.05, 15.90, 5.87 and 12.10%, for KNN, MLP, RFT and RLM, respectively. The superiority of RFT was also shown by the R2adj and Willmott's 'd'.

For example, the average R2adj for high and low-yield coffee was 0.866 and 0.868, respectively, being higher than that for the other tested algorithms. These results are similar to other authors, such as Meira, Rodrigues and Moraes (2008), who reached a precision of 73% in the prediction of infection rates of coffee rust at high-yields in the region of SO_{MG}.

The Willmott's 'd' indexes for high-yield coffee predicting with RFT were 0.966, 0.969, 0.969, 0.979, 0.970, 0.968, 0.973 for the ARG, ARX, BOE, CDM, MUZ, PAT, and VAR locations, respectively. For MLP, the Willmott's 'd' indexes for high-yield plants were 0.847, 0.715, 0.736, 0.392, 0.535, 0.797 and 0.687 for the ARG, ARX, BOE, CDM, MUZ, PAT and VAR locations, respectively. Pinto et al. (2002) also compared RLM and MLP in predicting coffee rust using as independent data the climatic variables in the SO_{MG} and observed that the smallest errors occurred with the use of MLP algorithms.

The prediction of coffee rust has already been studied by Pinto *et al.* (2002) using artificial neural networks, and by Meira, Rodrigues and Moraes (2008) using decision trees. Another technique was used by Girolamo Neto *et al.* (2014) using data mining techniques to develop coffee rust prediction models for the SO_{MG}, and they observed that for years of high and low-yields the best accuracies were 85.3% and 88.9%, respectively.

TABLE 8. Statistical indices of the algorithms in the prediction of coffee rust in high and low-yield coffee for the State of Minas Gerais, Brazil. Legend: BOE is Boa Esperança; ARG is Araguari; CDM is Carmo de Minas; ARX is Araxá; MUZ is Muzambinho; PAT is Patrocínio, VAR is Varginha, KNN is KNeighbors Regressor, MLP is Multi-layer Perceptron, RFT is Random Forest Regressor and RLM is Multiple Linear Regression (Ridge).

Locations	Yield	KNN	MLP	RFT	RLM	KNN	MLP	RFT	RLM	KNN	MLP	RFT	RLM
		R^2adj				RMSE				Willmott's 'd'			
ARG	High yield	0.719	0.634	0.864	0.833	19.837	21.213	8.807	14.091	0.821	0.847	0.966	0.932
ARX		0.727	0.504	0.868	0.790	19.994	22.865	8.906	14.932	0.823	0.715	0.969	0.916
BOE		0.705	0.619	0.860	0.755	19.670	21.779	8.353	16.399	0.800	0.736	0.969	0.867
CDM		0.763	0.688	0.874	0.823	19.723	32.473	8.198	16.106	0.849	0.392	0.979	0.915
MUZ		0.670	0.377	0.866	0.739	21.235	27.375	8.733	17.692	0.776	0.535	0.970	0.864
PAT		0.768	0.644	0.859	0.849	14.970	16.872	7.369	11.232	0.836	0.797	0.968	0.935
VAR		0.736	0.615	0.871	0.798	21.709	26.497	9.270	17.937	0.838	0.687	0.973	0.897
	Average	0.727	0.583	0.866	0.798	19.591	24.153	8.519	15.484	0.820	0.673	0.971	0.904
ARG	Low yield	0.754	0.647	0.867	0.854	17.457	19.454	7.818	12.029	0.854	0.833	0.971	0.942
ARX		0.703	0.477	0.871	0.800	18.540	21.978	8.265	13.785	0.803	0.656	0.966	0.921
BOE		0.684	0.441	0.871	0.737	7.117	8.918	2.557	5.654	0.782	0.555	0.975	0.859
CDM		0.688	0.502	0.858	0.823	8.783	10.489	3.806	16.106	0.785	0.650	0.965	0.915
MUZ		0.481	0.334	0.870	0.612	19.341	21.121	7.418	15.963	0.609	0.495	0.960	0.773
PAT		0.743	0.657	0.865	0.844	15.126	15.897	6.786	10.601	0.827	0.836	0.970	0.943
VAR		0.668	0.541	0.877	0.701	12.027	13.497	4.464	10.570	0.792	0.659	0.975	0.831
	Average	0.674	0.514	0.868	0.767	14.056	15.908	5.873	12.101	0.779	0.669	0.969	0.883

The RFT algorithm showed the highest accuracy (RMSE) and the MLP the lowest performance for both fruit loads in the cercospora forecast (Table 9).

The Willmott's 'd' indexes for the high-yield coffee predicting with RFT were 0.939, 0.887, 0.928, 0.840, 0.654, 0.865 and 0.637 for the ARG, ARX, BOE, CDM, MUZ, PAT and VAR locations, respectively. The Willmott's 'd' indexes for the MLP at high-yields were 0.644, 0.487, 0.628, 0.640, 0.654, 0.565 and 0.637 for the ARG, ARX, BOE, CDM, MUZ, PAT and VAR locations, respectively. The mean accuracy difference between these algorithms was 0.259% (Table 9).

The average R^2adj for the prediction of the cercospora in high-yield coffee were of 0.667, 0.507, 0.736 and 0.655 for the algorithms KNN, MLP, RFT and RLM, respectively. At low-yields the R2adj values were 0.632, 0.525, 0.725 and 0.628 for KNN, MLP, RFT and RLM, respectively. The authors Souza et al. (2013a), who used the decision tree methodology to evaluate cercosporiosis in conventional and organic coffee plantations, observed that the calibrated models had a 60% accuracy in the SO_{MG}.

TABLE 9. Statistical indices of the algorithms in the forecast of the cercospora in high and low-yield for the State of Minas Gerais, Brazil. Legend: BOE is Boa Esperança; ARG is Araguari; CDM is Carmo de Minas; ARX is Araxá; MUZ is Muzambinho; PAT is Patrocínio, VAR is Varginha, KNN is KNeighbors Regressor, MLP is Multi-layer Perceptron, RFT is Random Forest Regressor and RLM is Multiple Linear Regression (Ridge).

Locations	Yield	KNN	MLP	RFT	RLM	KNN	MLP	RFT	RLM	KNN	MLP	RFT	RLM
		R2adj				RMSE				Willmott's 'd'			
ARG	High yield	0.726	0.495	0.895	0.763	20.337	27.003	7.298	16.019	0.823	0.644	0.939	0.913
ARX		0.665	0.386	0.786	0.573	12.387	16.741	6.089	13.178	0.763	0.487	0.887	0.710
BOE		0.594	0.540	0.840	0.536	4.057	4.249	4.491	4.049	0.689	0.628	0.928	0.741
CDM		0.688	0.529	0.529	0.664	2.849	3.316	3.158	2.606	0.774	0.640	0.840	0.817
MUZ		0.587	0.611	0.611	0.718	6.304	6.295	3.295	5.025	0.698	0.654	0.654	0.845
PAT		0.716	0.421	0.921	0.621	5.731	7.035	7.035	5.515	0.775	0.565	0.865	0.830
VAR		0.689	0.569	0.569	0.707	2.804	3.192	3.192	2.330	0.792	0.637	0.637	0.843
	Average	0.666	0.507	0.736	0.655	7.781	9.690	4.937	6.960	0.759	0.608	0.821	0.814
ARG	Low yield	0.725	0.500	0.500	0.753	19.291	26.608	4.847	16.061	0.824	0.651	0.651	0.909
ARX		0.561	0.433	0.733	0.503	9.678	13.133	7.196	9.709	0.706	0.478	0.978	0.673
BOE		0.668	0.536	0.936	0.597	3.106	3.518	3.518	3.095	0.724	0.553	0.853	0.766
CDM		0.526	0.541	0.841	0.614	2.615	2.591	2.591	2.270	0.641	0.618	0.818	0.788
MUZ		0.560	0.516	0.516	0.632	5.597	5.832	5.832	5.049	0.673	0.561	0.961	0.792
PAT		0.788	0.621	0.621	0.743	2.972	3.270	6.971	2.325	0.846	0.724	0.773	0.913
VAR		0.598	0.533	0.933	0.554	1.780	1.905	1.905	1.648	0.701	0.562	0.762	0.742
	Average	0.632	0.525	0.725	0.628	6.434	8.122	4.694	5.737	0.731	0.592	0.828	0.798

In predicting the severity of the coffee bean miner using climate variables, the RFT algorithm showed the highest accuracy (RMSE) and the highest precision (R^2adj) in both fruit loads and in all localities, while MLP had the lowest accuracy (Table 10).

In the prediction of the coffee miner in high-yield plants, the Willmott's 'd' of the RFT were 0.864, 0.834, 0.842, 0.848, 0.845, 0.829 and 0.849 for the ARG, ARX, BOE, CDM, MUZ, PAT and VAR, respectively. The MLP for high-yield coffee showed values of 0.439, 0.570, 0.598, 0.741, 0.806, 0.447 and 0.886 for the ARG, ARX, BOE, CDM, MUZ, PAT and VAR locations, respectively. The mean accuracy difference between these algorithms was 0.240% (Table 10).

The average R^2adj for the prediction of the coffee miner in high-yield plants was 0.615, 0.486, 0.850 and 0.643 for the KNN, MLP, RFT and RLM algorithms, respectively. At low-yields the values of R2adj were 0.629, 0.471, 0.857 and 0.589, for KNN, MLP, RFT and RLM, respectively (Table 10). This demonstrates that the algorithms do not present prediction differences for the coffee miner bug at high or low-yields.

TABLE 10. Statistical indexes of the algorithms in the prediction of Coffee miner in high and low-yield crops for the State of Minas Gerais, Brazil. Legend: BOE is Boa Esperança; ARG is Araguari; CDM is Carmo de Minas; ARX is Araxá; MUZ is Muzambinho; PAT is Patrocínio, VAR is Varginha, KNN is KNeighbors Regressor, MLP is Multi-layer Perceptron, RFT is Random Forest Regressor and RLM is Multiple Linear Regression (Ridge).

Locations	Yield	KNN	MLP	RFT	RLM	KNN	MLP	RFT	RLM	KNN	MLP	RFT	RLM
		R2adj				RMSE				Willmott's 'd'			
ARG	High yield	0.688	0.197	0.858	0.656	14.367	22.430	6.645	15.483	0.789	0.439	0.864	0.832
ARX		0.426	0.537	0.839	0.602	4.982	5.438	1.063	4.126	0.622	0.570	0.834	0.716
BOE		0.426	0.516	0.840	0.530	4.982	2.147	1.018	2.033	0.622	0.598	0.842	0.689
CDM		0.673	0.655	0.859	0.575	1.339	1.372	0.662	1.365	0.766	0.741	0.848	0.727
MUZ		0.708	0.723	0.856	0.643	2.167	2.108	1.201	2.261	0.780	0.806	0.845	0.779
PAT		0.674	0.049	0.842	0.643	11.282	14.375	4.061	10.399	0.709	0.447	0.829	0.836
VAR		0.709	0.824	0.855	0.854	1.087	0.880	0.565	12.029	0.799	0.886	0.849	0.942
	Average	0.615	0.486	0.850	0.643	5.744	6.964	2.174	6.814	0.727	0.641	0.844	0.789
ARG	Low yield	0.723	0.454	0.850	0.664	14.505	19.605	7.433	14.776	0.801	0.675	0.858	0.844
ARX		0.478	0.535	0.861	0.554	8.148	8.225	3.716	7.591	0.627	0.573	0.842	0.668
BOE		0.622	0.403	0.848	0.467	2.457	2.871	0.953	2.227	0.723	0.457	0.850	0.623
CDM		0.659	0.545	0.863	0.625	1.655	1.854	0.775	1.601	0.749	0.619	0.854	0.762
MUZ		0.681	0.537	0.852	0.621	2.539	2.946	1.225	2.292	0.776	0.590	0.842	0.751
PAT		0.599	0.140	0.868	0.628	9.537	11.117	4.416	8.387	0.690	0.642	0.851	0.881
VAR		0.638	0.684	0.859	0.564	1.273	1.224	0.588	1.232	0.735	0.747	0.850	0.712
	Average	0.629	0.471	0.857	0.589	5.731	6.835	2.729	5.444	0.729	0.615	0.849	0.749

In the prediction of the coffee borer, the RFT algorithm showed the highest accuracy in both productive loads and in all localities, whereas the KNN showed the lowest accuracy in the predictions (Table 11).

The indices of Willmott's 'd' of the RFT to predict the coffee borer at high-yields were of 0.955, 0.933, 0.969, 0.930, 0.955, 0.954 and 0.963 for the localities of ARG, ARX, BOE, CDM, MUZ, PAT and VAR, respectively. However, the Willmott's 'd' indices for KNN were 0.721, 0.740, 0.813, 0.643, 0.823, 0.694 and 0.793 for the ARG, ARX, BOE, CDM, MUZ, PAT and VAR, respectively. The mean accuracy difference between these algorithms was 0.214% (Table 11).

The average R_2adj for the prediction of coffee borer in high-yield plants was 0.633, 0.672, 0.854 and 0.744 for the KNN, MLP, RFT and RLM algorithms, respectively. At low-yields the R2adj were 0.633, 0.672, 0.854 and 0.744, for KNN, MLP, RFT and RLM, respectively (Table 11). This demonstrates that the algorithms do not present differences in prediction of coffee borer at high or low-yields.

TABLE 11. Statistical indices of the algorithms in the prediction of coffee borer in high and low-yield coffee for the State of Minas Gerais, Brazil. Legend: BOE is Boa Esperança; ARG is Araguari; CDM is Carmo de Minas; ARX is Araxá; MUZ is Muzambinho; PAT is Patrocínio, VAR is Varginha, KNN is KNeighbors Regressor, MLP is Multi-layer Perceptron, RFT is Random Forest Regressor and RLM is Multiple Linear Regression (Ridge).

Locations	Yield	KNN	MLP	RFT	RLM	KNN	MLP	RFT	RLM	KNN	MLP	RFT	RLM
		R2adj				RMSE				Willmott's 'd'			
ARG	High yield	0.621	0.501	0.855	0.778	4.870	5.369	2.229	5.621	0.721	0.604	0.955	0.882
ARX		0.635	0.512	0.852	0.730	4.004	4.489	2.223	4.307	0.740	0.549	0.933	0.841
BOE		0.692	0.857	0.865	0.811	1.295	0.936	0.550	1.019	0.813	0.901	0.969	0.889
CDM		0.500	0.629	0.838	0.694	2.197	2.150	0.853	1.387	0.643	0.519	0.930	0.802
MUZ		0.713	0.838	0.851	0.740	0.923	0.736	0.476	0.855	0.823	0.887	0.955	0.843
PAT		0.592	0.545	0.858	0.834	2.155	2.244	0.919	1.521	0.694	0.627	0.954	0.912
VAR		0.679	0.825	0.861	0.618	0.557	0.433	0.227	0.519	0.793	0.887	0.963	0.747
	Average	0.633	0.672	0.854	0.744	2.386	2.337	1.068	2.175	0.747	0.751	0.951	0.845
ARG	Low yield	0.654	0.641	0.862	0.770	4.356	4.442	1.815	5.402	0.752	0.726	0.969	0.874
ARX		0.596	0.864	0.877	0.671	2.677	1.873	1.075	2.597	0.680	0.873	0.951	0.806
BOE		0.593	0.902	0.857	0.797	0.987	0.541	0.395	0.658	0.723	0.938	0.957	0.878
CDM		0.542	0.950	0.830	0.689	1.376	0.532	0.827	1.025	0.661	0.969	0.937	0.799
MUZ		0.619	0.621	0.853	0.735	1.339	1.338	0.444	0.861	0.738	0.720	0.963	0.839
PAT		0.715	0.859	0.835	0.752	1.974	1.485	1.159	1.906	0.805	0.903	0.939	0.867
VAR		0.531	0.872	0.862	0.574	0.410	0.238	0.147	0.385	0.695	0.916	0.971	0.704
	Average	0.607	0.815	0.854	0.712	1.874	1.493	0.837	1.833	0.722	0.864	0.955	0.824

The RMSE maps of the RFT for the SO_{MG} and CE_{MG} regions for rust, cercospora, coffee miner and coffee borer can be observed in Figures 12-15.

The spatial variability of the RMSE index for the RFT algorithm shows the accuracy of this model in the prediction of rust in high-yield (Figure 12.A) and low-yield (Figure 12.B) coffee, using the climatic elements as independent variables. The RFT in the prediction of Rust in high-yield coffee demonstrated a more constant variability of RMSE in the SO_{MG} and CE_{MG} regions. Patrocínio had an RMSE of only 7.3% (Figure 12.A). In the prediction of rust in low-yield coffee the SO_{MG} region showed the lowest RMSE value of 2.55% (BOE), while in the CE_{MG} region the values reached 8.26% (ARX). RMSE of only 2.55% is considered very accurate in prediction models using climatic data (Moreto and Rolim, 2015, Marca *et al.,* 2015).

The RFT in predicting the severity indexes of cercospora showed greater accuracy in the SO_{MG} region in both crop loads. In SO_{MG} the RMSE values ranged from 3.19 to 4.49 for high-yield coffee (Figure 13.A) and between 1.90 to 5.83 for low-yield (Figure 13.B). The highest RMSE were found in the Araguari locality, reaching values of 7.29% in high-yield plants.

In the prediction of the Coffee miner the RFT algorithm demonstrated low RMSE values in both productive loads. In the SO_{MG} region the RMSE ranged from 0.56 to 1.20 for high-yield and between 0.587 and 1.22 for low-yield coffee (Figure 14). In Araguari the RFT demonstrated the highest values of RMSE in the prediction of the Coffee miner, reaching 7.43% in low-yield coffee.

The RFT algorithm demonstrated greater accuracy in the SO_{MG} region in both the productive loads in the prediction of the coffee borer. In SO_{MG} values ranged from 0.227 to 0.853 for high-yield and 0.147 and 0.827 for low-yield coffee. In the CE_{MG} the RMSE values were elevated. For example, in the prediction of coffee borer in Araguari the RMSE was of 2.229 and 1.815 in high and low-yield coffee, respectively (Figure 15).

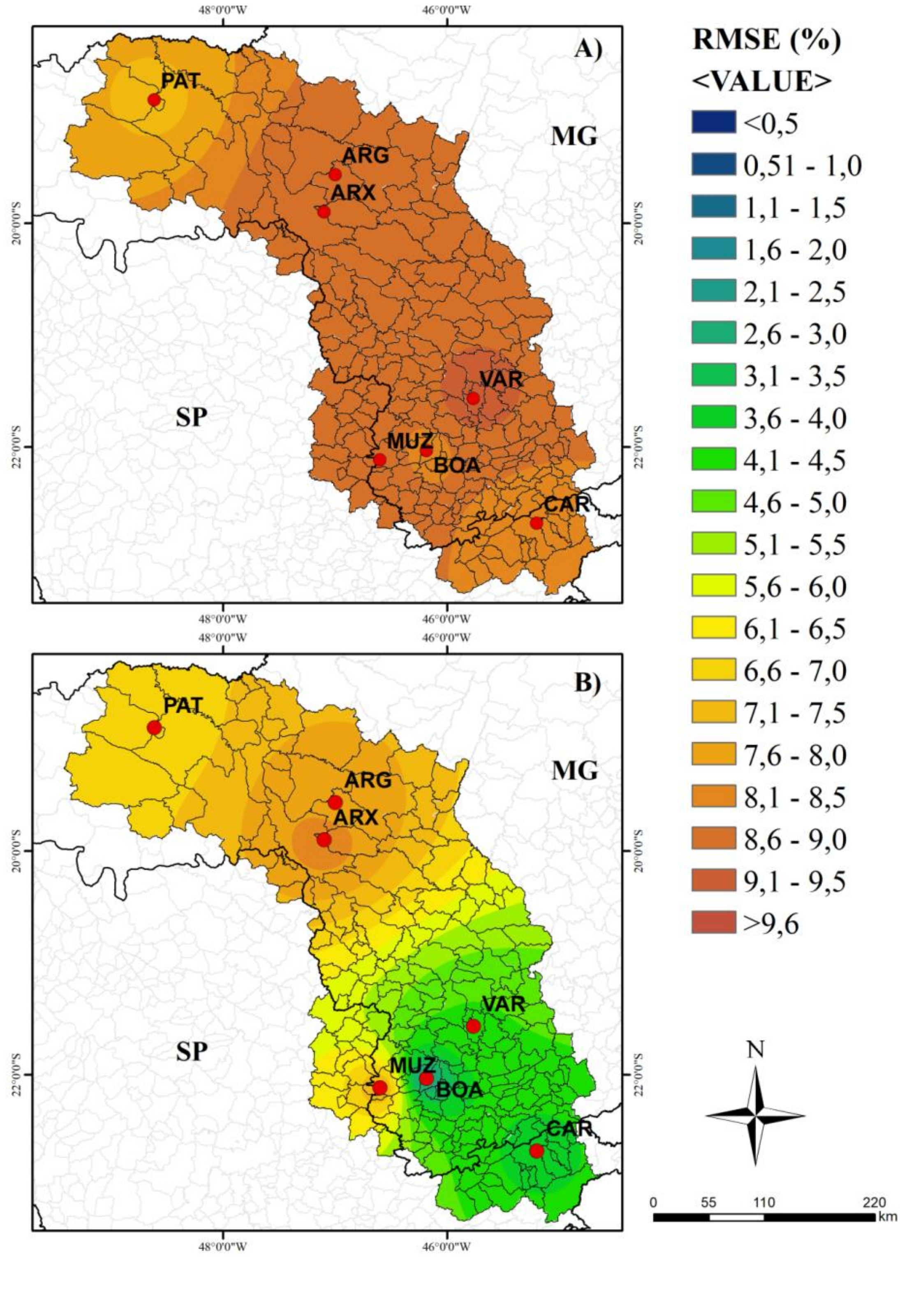

FIGURE 12. Spatialization of the RMSE index (%) of the forecast of coffee rust at high (A) and low-yields (B) by the Random Forest Regressor algorithm for the south of Minas and Cerrado Mineiro regions, Brazil.

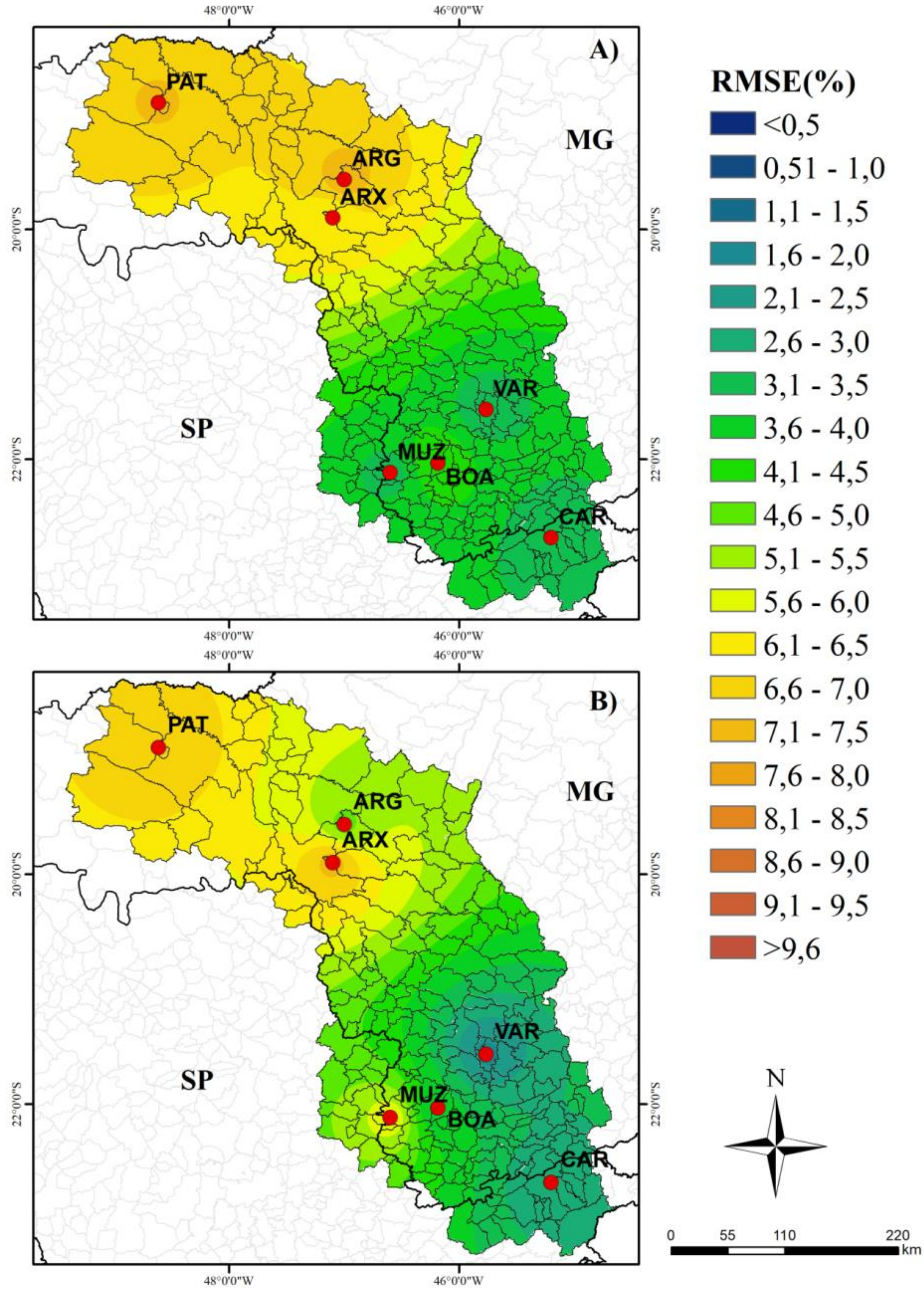

FIGURE 13. Spatialization of the RMSE index (%) of the forecast of the Cercospora at high (A) and low-yields (B) by the Random Forest Regressor algorithm for the south of Minas and Cerrado Mineiro regions, Brazil.

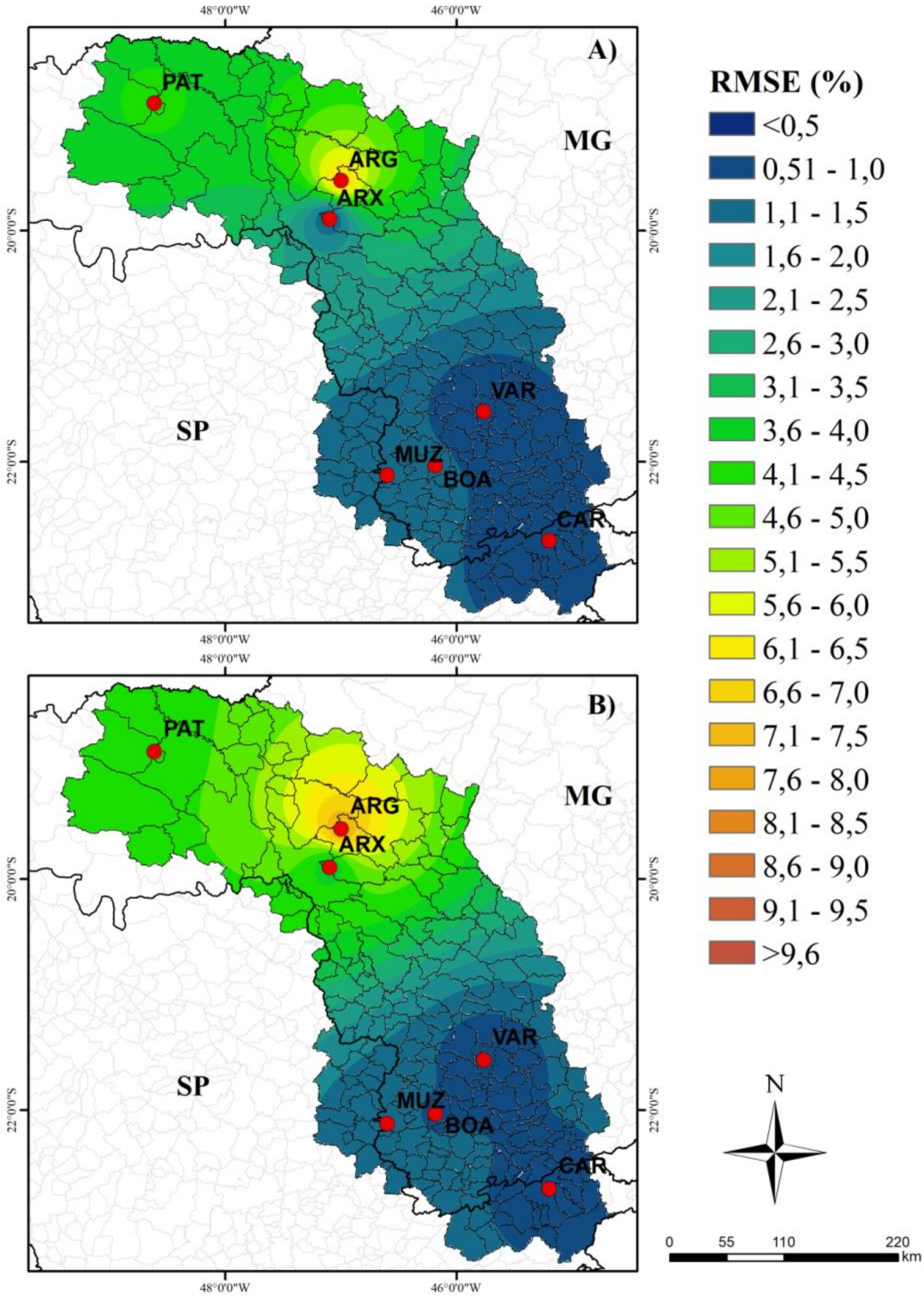

FIGURE 14. Spatialization of the RMSE index (%) of the forecast of the Coffee miner at high (A) and low-yields (B) by the Random Forest Regressor algorithm for the south of Minas and Cerrado Mineiro regions, Brazil.

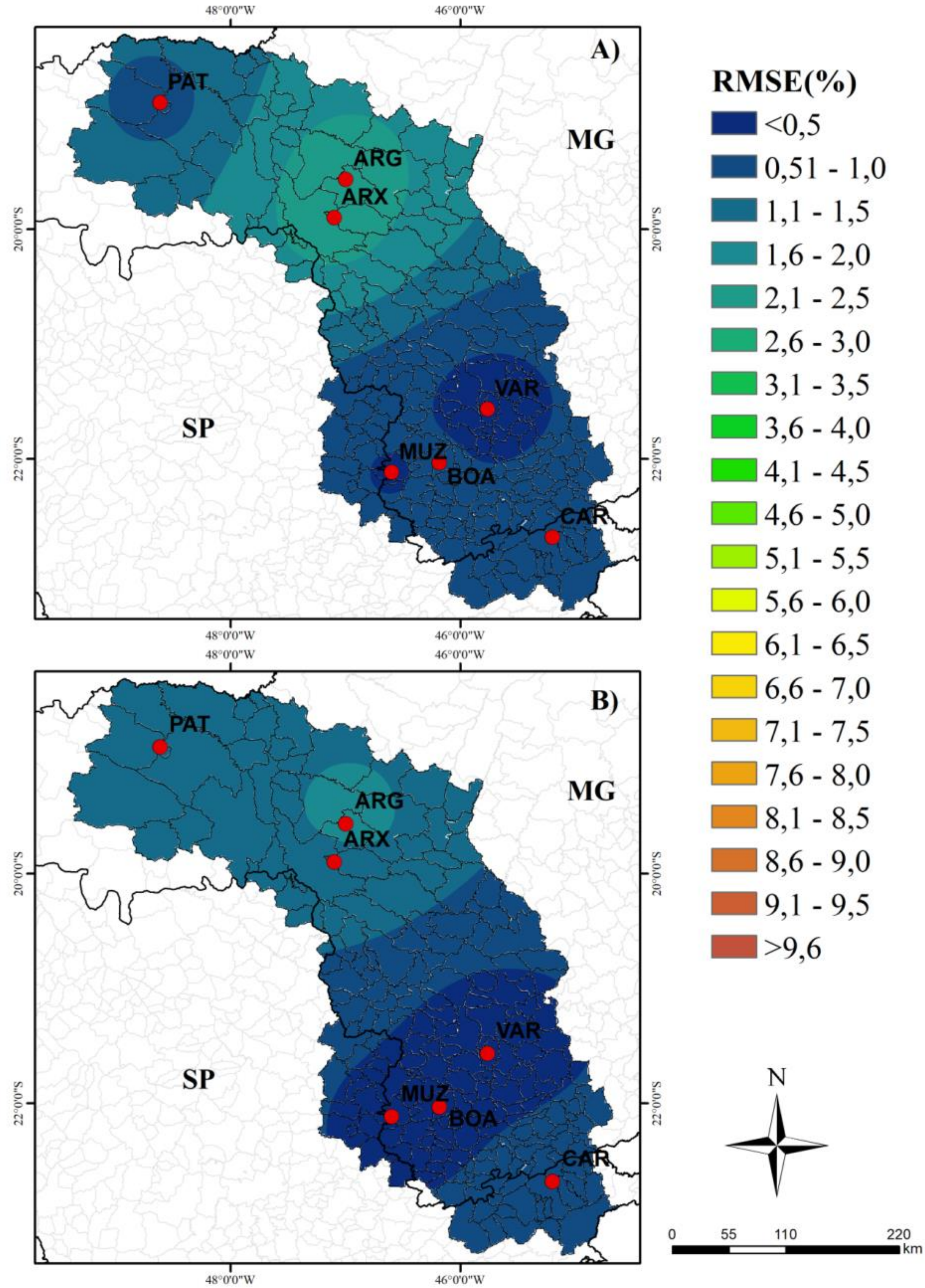

FIGURE 15. Spatialization of the RMSE index (%) of the forecast of the Coffee berry borer at high (A) and low-yields (B) by the Random Forest Regressor algorithm for the south of Minas and Cerrado Mineiro regions, Brazil.

Conclusions

The rust, cercospora, and miner showed a sigmoid progression curve, while the coffee borer showed a Gaussian trend in both crop fruit loads.

The maximum air temperature, the number of days with relative humidity above 80% and the relative humidity were the climatic variables that showed significant correlations with the coffee rust disease. Precipitation showed negative correlations with the coffee miner in both loads, which causes a decrease in the severity of the pest with increasing rainfall.

The predictive models of coffee diseases and pests developed in this work provide better subsidies for the monitoring of diseases in years of high and low-yield fruits.

The algorithms of machine learning do not present differences in the predictions between high and low-yield harvests. The random forest model was more accurate in the prediction of coffee rust, cercospora, coffee miner and coffee borer using climatic conditions. In the prediction of coffee rust, cercospora, and coffee miner, the neural networks presented the lowest performance, whereas for the coffee borer the k-Nearest neighbor's algorithm had the lowest performance.

Acknowledgements

This research was supported by the São Paulo Research Foundation (FAPESP, 2015/17797-4).

CAPÍTULO 4 – Validation of ERA-Interim (ECMWF) surface climatic data and implications for modelling water balance

ABSTRACT - Gridded meteorological systems greatly facilitate the analysis of the impacts of climate on crop development and productivity. Comparisons of these data with actual surface data validate this data source for various analyses in agricultural areas. The impact of the use of these grid data is an important evaluation for the temporal and spatial simulation of soil-water availability for crops. We thus sought to determine the accuracy of climatic data from the European Centre for Medium Range Weather Forecast (ECMWF) with the meteorological ground stations, and tested its application for modelling climatic water balance. Monthly data for air temperature (T) and precipitation (P) from ECMWF were compared with the data from 771 surface stations (National Meteorological Institute, INMET) in the state of Minas Gerais in southeastern Brazil for 1979-2017. Potential evapotranspiration was estimated by the Thornthwaite method (1948), and water balance was estimated by the method proposed by Thornthwaite and Mather (1955), with an available water capacity of 100 mm. We temporally and spatially compared the two data sources, and the comparisons were evaluated for accuracy using mean absolute percentage error (MAPE) root mean square error (RMSE) and for precision using the adjusted coefficient of determination (R^2adj). ECMWF T and P tended to be temporally and spatially similar to the INMET data, with R^2adj = 0.95±0.0017, MAPE = 13.08±23.39%, and RMSE = 0.86±0.42 °C for T and R^2adj = 0.95±0.003, MAPE = 14.10±17.01%, and RMSE = 1.64±0.58 mm for P. The largest deviation between INMET T and ECMWF T was 2.81 °C, mainly in the southwest of the state (the Minas Gerais triangle) and part of the central region during winter and spring, and the smallest deviation was -0.19 °C in the northeast. The largest deviation between INMET P and ECMWF P was 75 mm mo^{-1} in the summer, mainly between January and February in the central region of Minas Gerais. ECMWF T and ECMWF P allowed an accurate estimation of the components of the water balance. For example, the lowest MAPEs were 1.21% for ECMWF water-storage capacity (southern Minas Gerais), 9.16% for ECMWF water deficiency (Vale do Jequitinhonha e Mucurí), and 8.69% for ECMWF excess water (Vale do Jequitinhonha e Mucurí). The average deviations were ±10 mm mo^{-1} between the INMET and ECMWF water-storage capacities, ±7.6 mm mo^{-1} between the INMET and ECMWF water deficiencies, and ±23.6 mm mo^{-1} between the INMET and ECMWF water excesses. We concluded that the climatic variables from the ECMWF system were more accurate and could be used to model the climatological water balance.

KEY-WORDS: Climatic zoning; forecast verification; water deficiency; climatic variables; general circulation model.

Introduction

Agricultural production is greatly affected by climatic variability and extreme events (Ceglar *et al*., 2016), and the accuracy of models forecasting regional productivity is strongly linked to the sources of meteorological and soil data (Orth *et al*., 2016). Data sources may be limited by temporal data failures, calibrations, possible sensor problems, and spatial scales. Countries and even regions with few surface meteorological stations (Rodrigues, Moretto, Guilhoto, 2015) or with data gaps (Pereira *et al*., 2002) may benefit from gridded global data (GGD) systems to resolve problems where meteorological conditions are important. The European Centre for Medium-Range Weather Forecast (ECMWF) is a GGD system that provides free and relatively simple grid data.

ECMWF collects information from a variety of meteorological sources around the world, such as meteorological and satellite radar systems (ECMWF, 2009; Couto *et al*., 2015), and is an important source of data. ECMWF data are processed on a 10-day period scale at a spatial resolution of 27 × 27 km, freely available for downloading from the Joint Research Centre meteorological database of the European Commission (Grazziani, 2007; Moraes; Rocha; Lamparelli, 2014) that processes information for rainfall, air temperature, global solar radiation, and evapotranspiration using the Penman-Monteith equation.

The surface of the Earth is an essential component of the climatic system (Orth *et al*., 2016), interacting with the atmosphere by the exchange of water and energy, and can also accumulate and maintain anomalies induced by atmospheric forcing (Aquila *et al*., 2016). Surface meteorological stations thus record micrometeorological conditions, and GGD systems provide data for meso to micrometeorological conditions.

The use of meteorological information from the ECMWF atmospheric system is an alternative to the use of surface meteorological data. Few studies, however, have compared the accuracies of the ECMWF and surface data. Some studies have evaluated published global models. For example, Sodoudi *et al*. (2010) compared the accuracy of ECMWF precipitation data with surface-station data in Iran and concluded that the ECMWF system performed better for high mountainous regions than for flat terrain and deserts, with an error of 0.46 mm d^{-1}. McDonnell *et al*. (2018) verified and corrected the ECMWF prediction error for Irish meteorological stations and found that the ECMWF system performed well, with a mean RMSE of 5.56 mm d^{-1}. These data from atmospheric models can also be used to quantify soil water, but this application has not yet been tested.

Agriculture is one of the human activities most affected by climatic conditions (Sá Júnior *et al*., 2012), and the availability of soil water is the main cause of variation in crop yields (Martinez *et al*. 2013, Schrader *et al*., 2013). Cultivars can only use soil water within the reach of their root systems (Moreira *et al*., 2014), and productivity is substantially reduced due to the coincidence of periods with low rainfall during the growing season (Faria and Bowen, 2003). For example, the timing of crop establishment and flowering presents a high risk, because frequent water deficits in short dry periods associated with high evapotranspiration strongly affect the development and productivity of agricultural crops.

Quantifying reductions in productivity as a function of water deficits requires simulating the components of the water balance (WB) of the soil (Negm *et al*., 2013). Many conventional stations are required for precise simulations, and the data must be accurate, but these conditions have not yet been met. Using grid data for simulating WBs is one method to address errors and lack of data from surface stations.

WB is an account of the amount of water in the soil (Brunel-Saldias *et al*., 2018) from the application of the principle of mass storage, which states that water storage (STO) is the result of inputs and outflows of water in a volume of soil (Moreira *et al*., 2014; Abatzoglou *et al*., 2018). Soil STO variability is calculated as:

$$\Delta STO = \left(P + I + DE + RU + DLi + CA\right) - \left(ET + RO + DLo + DR\right) \qquad (1)$$

where ΔSTO is the variation in storage over time; precipitation (P), irrigation (I), dew (DE), run in (RU), the input of lateral drainage (DLi), and capillary ascension (CA) runoff are the water-input components in the soil profile; and evapotranspiration (ET), runoff (RO), the outflow of lateral drainage (DLo), and drainage (DR) to the water table are the water-outflow components. All components or flows are given in mm $time^{-1}$, and the number and form of simulations of these components indicate whether a WB model can be classified as empirical (functional) or mechanistic.

WB models were first developed and applied for climatological purposes in the 1940s (Thornthwaite, 1948) and 1950s (Thornthwaite and Mather, 1955). These WB models simulated the balance between inputs from rainwater and ice melt and outflows from evapotranspiration, ice flow, and deep drainage. They have become universal due to their simplicity and ease of use. The Thornthwaite and Mather (1955) (WBTM) model was quickly adopted for irrigation and management purposes.

Several studies have applied the WBTM model in its original form (Alley, 1984). For example, Pereira (1986) successfully used the original WBTM model to study the minimum and maximum water storage of a podzolic soil in the state of São Paulo, Brazil. Other WB models such as BUD (Budyko, 1958), FAO (Doorenbos and Pruitt, 1977), SWAP (Van-Dam, 1997), DSSAT-SWBM (Ritchie, 1998), and SWB (Dripps *et al*., 2003) have been proposed, but the WBTM model is very interesting due to its simplicity and high accuracy.

The use of ECMWF data instead of surface data to simulate regional WB components has been hypothesised, but we have found no studies comparing WB modelling with climatic data from surface stations and ECMWF data for Brazilian conditions. We thus sought to determine the accuracy of the climatic data from the ECMWF with the meteorological ground stations and tested its application for modelling climatic WBTM.

Material and methods

We used monthly meteorological data for mean air temperature (T) and rainfall (P) for 1979-2017 in the state of Minas Gerais, southeastern Brazil. Minas Gerais has an area of 586 528 km^2 and is located between 13.94 and 22.50°S and 41.73 and 52.87°W (Figure 1). The Köppen and Geiger (1928) climatic classification lists five classes (Am, Aw, BSh, Cwa, and Cwb) but Aw is the predominant class in Minas Gerais.

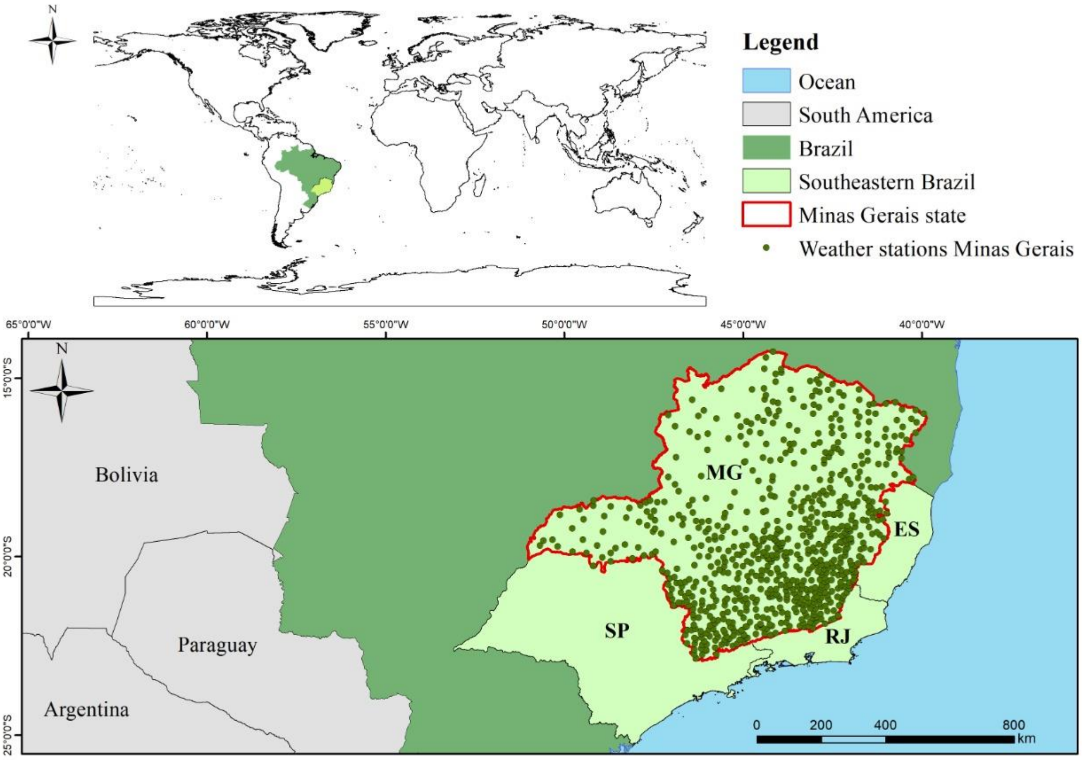

FIGURE 1. Location of the state of Minas Gerais (MG) in Brazil.

Data from ECMWF and meteorological stations of the National Institute of Meteorology (INMET) (Figure 2B) were used. Data were compiled from 771 INMET stations (Figure 2A) and 1578 points corresponding to ECMWF virtual stations (Figure 2C).

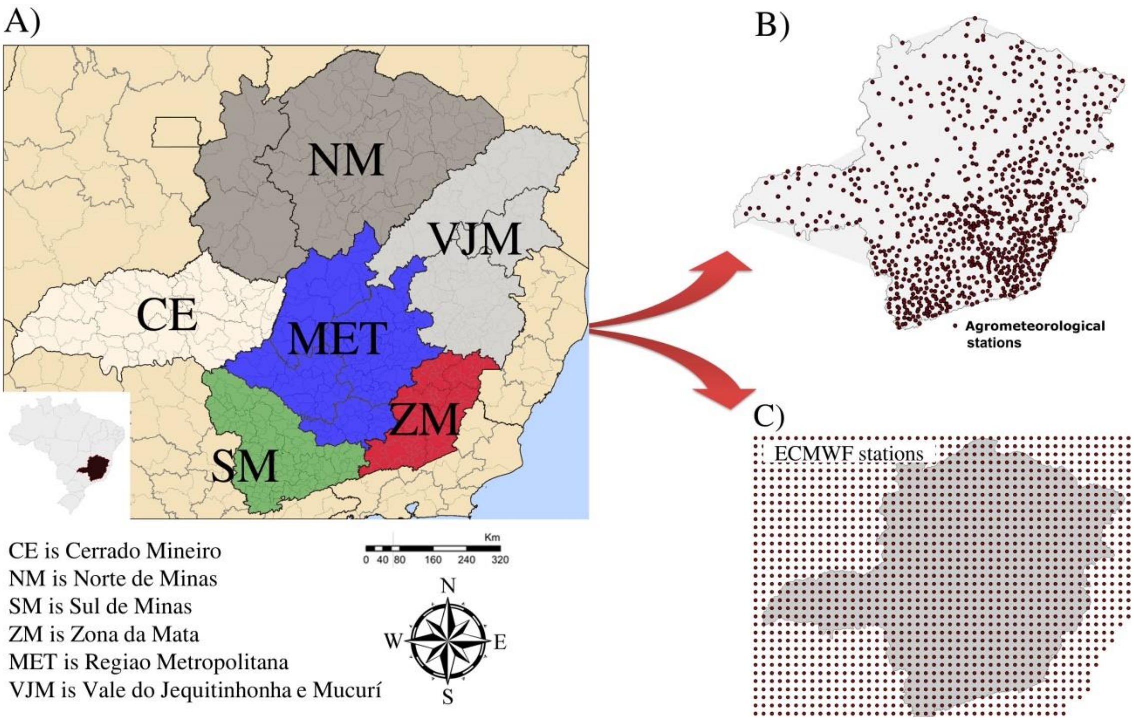

FIGURE 2. Regions of Minas Gerais (A) and locations of the INMET meteorological stations (B) and the ECMWF virtual stations (C).

Monthly ECMWF data were obtained, with a spatial resolution of 1°, and then pre-processed and transformed into 0.25° (±25 × 25 km) grids. Data from the ECMWF system were freely collected at http://www.ecmwf.int/. The data were interpolated, and the meshes generated were overlapped to select the ECMWF points that corresponded to the geographic coordinates of the INMET stations. The coordinates and overlapping meshes allowed us to compare the two data sources. The comparison between the T and P data was stratified relative to the seasons and as a function of the regions of Minas Gerais. Summer, autumn, winter, and spring were defined as January-March, April-June, July-September, and October-December, respectively. This stratification was necessary to more accurately represent the peculiarities of the ECMWF system in each region and season.

Modelling Water Balance

Potential evapotranspiration (PET) was calculated using the method of Thornthwaite (1948):

$$PET = ETp \cdot COR \tag{2}$$

$$\begin{cases} ETp = 16 \cdot \left(10 \cdot \frac{Tn}{I}\right)^{a}; & when\ 0 \leq Tn < 26.5\ °C \\ ETp = -415.85 + 32.24 \cdot Tn - 0.43 \cdot Tn^{2}; & when\ Tn \geq 26.5\ °C \end{cases} \tag{3}$$

$$I = \sum_{n=1}^{12} (0.2 \cdot Tn)^{1.514} \tag{4}$$

$$a = 0.49239 + 1.7912 \cdot 10^{-2} \cdot I - 7.71 \cdot 10^{-5} \cdot I^{2} + 6.75 \cdot 10^{-7} \cdot I^{3} \tag{5}$$

$$COR = \frac{N}{12} \cdot \frac{NDP}{30} \tag{6}$$

where ETp is standard evapotranspiration, COR is a correction factor based on the actual number of days and the photoperiod of the month, Tn is the mean temperature of month n, I is an index for the heat level in the region, a is a regional thermal index, NDP is the number of days of the period in question, and N is the average photoperiod of the month in question.

PET data from INMET and ECMWF were used to generate the WB components. The WBTM model with an available water capacity of 100 mm was used:

$$if\ \ (P - PET)_i < 0 = \begin{cases} NAC_i = NAC_{i-1} + (P + PET)_i \\ STO_i = WCe^{\frac{(NAC_i)}{WC}} \end{cases} \tag{7}$$

$$if\ \ (P - PET)_i \geq 0 = \begin{cases} STO_i = (P - PET)_i + STO_{i-1} \\ NAC_i = WCln^{\frac{(STO_i)}{WC}} \end{cases} \tag{8}$$

$$ALT_i = STO_i - STO_{i-1} \tag{9}$$

$$AET_i = \begin{cases} P + |ALT_i| & , when\ ALT < 0 \\ PET_i & , when\ ALT \geq 0 \end{cases} \tag{10}$$

$$DEF = PET - AET \tag{11}$$

$$EXC_i = \begin{cases} 0 & , when\ WC < 0 \\ (P - PET)_i - ALT_i & , when\ WC = 0 \end{cases} \tag{12}$$

where PET is potential evapotranspiration (mm), WC is available water capacity (mm), STO is soil-water storage (mm), NAC is the difference between total rainfall and PET, P is rainfall (mm), DEF is water deficiency in the soil-plant-atmosphere system (mm), AET is actual evapotranspiration (mm), EXC is excess water in the soil-plant-atmosphere system (mm), ALT is the difference between STO of the current and the preceding month (mm), and *i* is the month.

After completing the WB calculations, the ECMWF data ($T_{ECMWF;}$ $P_{ECMWF;}$ $STO_{ECMWF;}$ $DEF_{ECMWF;}$ EXC_{ECMWF}) and the data of the surface meteorological stations ($T_{INMET;}$ $P_{INMET;}$ $STO_{INMET;}$ $DEF_{INMET;}$ EXC_{INMET}) were compared by statistical indicators (accuracy and precision). The accuracy or accuracy, that is, the estimate is close to the observed value, was evaluated by MAPE (Mean Absolute Error) and RMSE (root mean square error). Precision is the ability of the model to repeat the estimate and was evaluated by the R^2 (coefficient of determination) adjusted according to Cornell and Berger (1987) and the p-value of the regression (Table 1).

TABLE 1. Indices of statistical accuracy and precision used for evaluating the climatic data from ECMWF and INMET. Yest, estimated value of y; Yobs, observed value of y; Xobs, observed value of x; n, number of datapoints; and k, number of independent variables in the regression. Overlined variables indicate averages.

Statistical indexes	**Equations**
Accuracy	
MAPE	$MAPE = \frac{\sum_{i=1}^{N}\left(\left\lvert\frac{Yest_i - Yobs_i}{Yobs_i}\right\rvert \cdot 100\right)}{N}$
RMSE	$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(Y_{obs_i} - Y_{est_i}\right)^2}{N}}$
Precision	
R^2	$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(Yest_i - \bar{Y}obs\right)^2}{\sum_{i=1}^{n}\left(Yest_i - \bar{Y}obs\right)^2 - \sum_{i=1}^{n}\left(Yest_i - Yobs_i\right)^2}$
R^2adj	$R^2\ adjusted = \left[1 - \frac{(1 - R^2)\cdot(N-1)}{N-k-1}\right]$

Maps were generated using kriging interpolation (KRIGE, 1951) with 1 neighbour and a resolution of 0.25 ° in the spherical model (equation 13). The meshes and semivariograms were adjusted for determining the values of the nugget effect, structural variation (the difference between the level and the nugget effect), range for all months in exponential, spherical, and Gaussian models. These settings allowed a better comparison with the ECMWF grid data.

$$\hat{Z}(x) - m(x) = \sum_{i=1}^{n(x)} \lambda i\ (x)[Z(x_1) - m(xi)] \quad (13)$$

where: $\lambda i\ (x)$ is observation weights $Z(x_1)$; $Z(x_1)$ is interpreted as the realization of VAZ (x); VAZ (x) is Semivariogram modeling $m(x)$, is the expected value of Z (x) at the point x; $n(x)$, is the number of data inside a neighborhood x.

Results and discussion

Air temperature data

T_{ECMWF} indicated a temporal trend similar to the T_{INMET} data, with T highest from September to April (summer) and lowest from April to August (winter) in each region of Minas Gerais. This trend of variability was also reported by Alvares *et al*. (2013). Mean R^2adj, MAPE, and RMSE for the temporal variation in all periods and regions were 0.95±0.0017, 13.08±23.39%, and 0.86±0.42 °C, respectively. The details of R^2adj, MAPE, and RMSE and between the temporal T_{INMET} and T_{ECMWF} data for each region are shown in Figure 3. The temporal accuracy of the ECMWF system was highest for Zona da Mata (ZM) and Vale do Jequitinhonha e Mucurí (VJM), where mean MAPE and RMSE were <12% and 1 °C, respectively, for the entire year. For example, the average monthly T for ECMWF and INMET for VJM (Figure 3F), where MAPE was lowest (5.1%), was 24.5 and 24.7 °C, respectively.

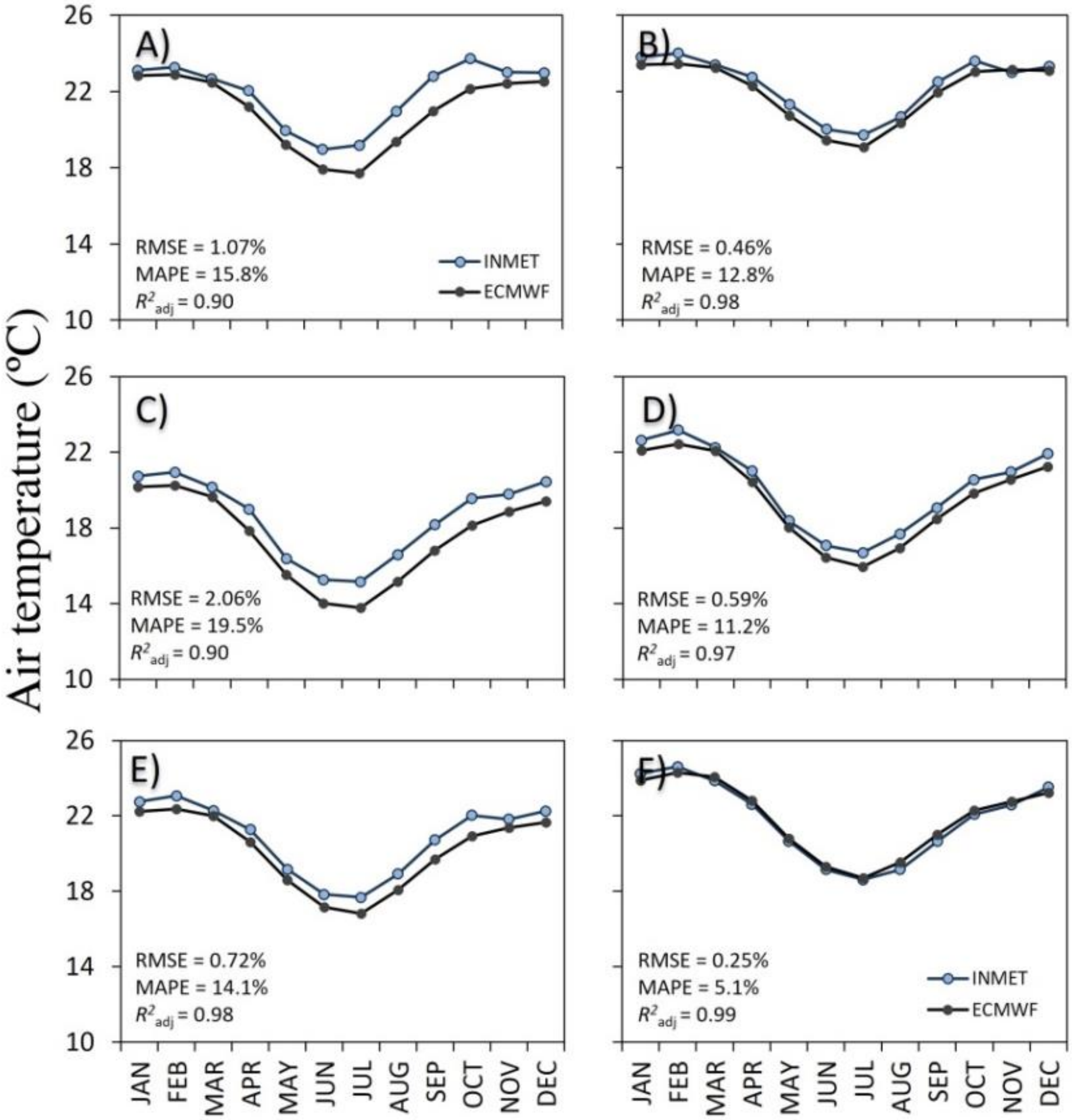

FIGURE 3. Temporal variation of the data for mean air temperature from INMET surface stations and ECMWF for the main regions of Minas Gerais, 1979-2017. A) Cerrado Mineiro, B) Norte de Minas, C) Sul de Minas, D) Zona da Mata, E) Regiao Metropolitana, and F) Vale do Jequitinhonha e Mucurí.

The spatialisation of T_{INMET} (Figure 4) and T_{ECMWF} (Figure 5) demonstrated that the ECMWF system could represent the spatial variability of T. T was lowest in the south and highest in the northeast, with means of 18.5 and 25.5 °C, respectively.

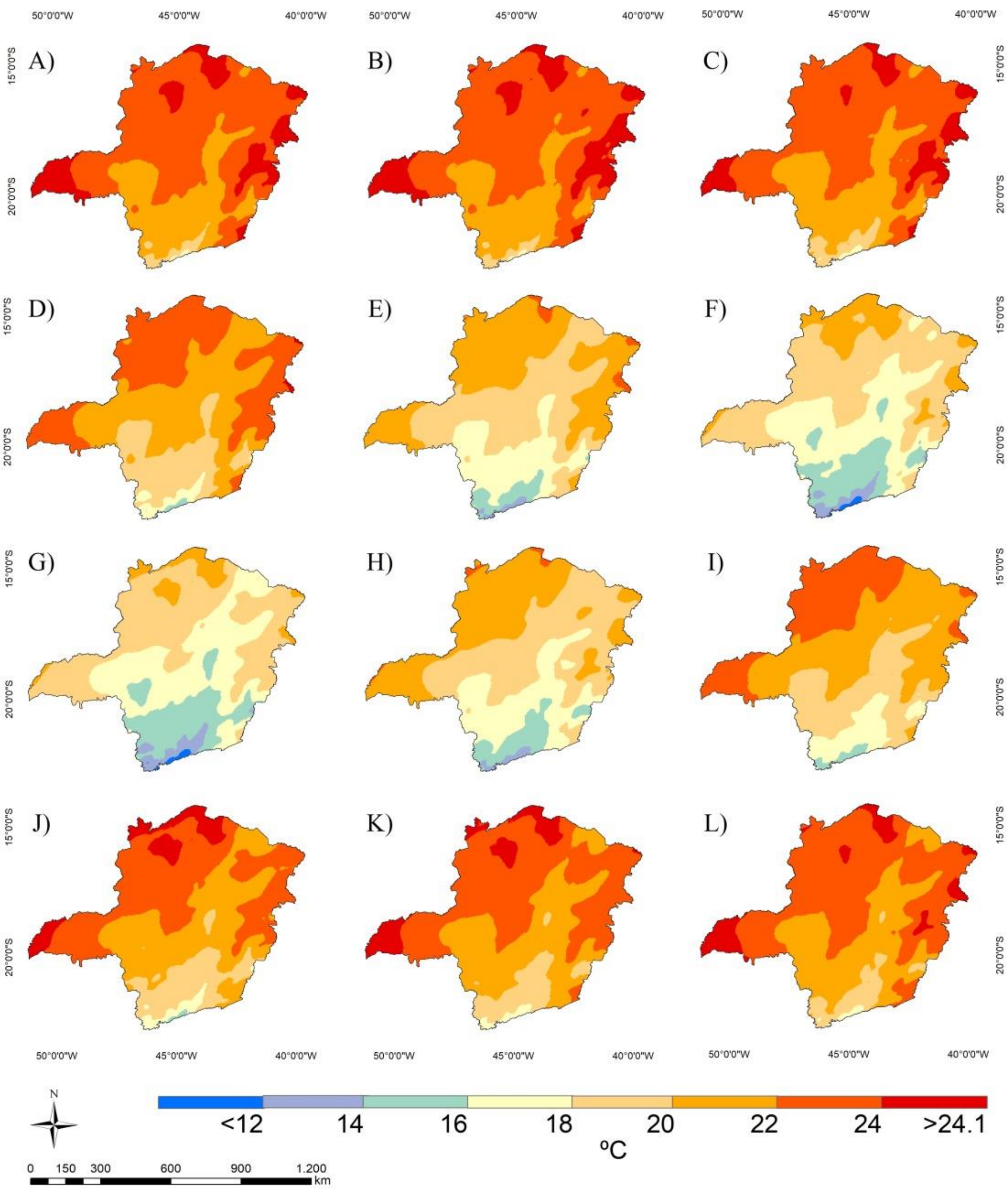

FIGURE 4. Spatial variation of the data for air temperature from INMET surface stations in Minas Gerais, 1979-2017. A) January, B) February, C) March, D) April, E) May, F) June, G) July, H) August, I) September, J) October, K) November, and L) December. Maps are from the spherical model with one neighbour and a resolution of 0.25°.

The nugget, sill, and range for the T_{INMET} data averaged 0.129±0.112, 1.14±0.139, and 127.61±14.71, respectively (Table 2). The sill of semivariogram adjustment is its limit (the sum of the nugget effect with the variation of the data) after stabilisation from a particular distance. The distance at which the semivariogram stabilises is known as the range. The sill of the semivariogram was lower in hot months, varying between 0.93 and 1.14 °C. The level of sill increased as T decreased, getting at 1.40 °C in June in the spherical model, indicating higher variance in the data pairs in months with lower Ts.

The nugget did not vary as a function of T, e.g. the pips effect averaged 0.12 and 0.13 °C in months with high and low Ts, respectively. The nugget is the semivariance for a distance of zero and represents the random variance that the semivariogram measures. Interestingly, the nugget is low because it represents measurement errors (Landim, 2003) or the general spatial dependence.

TABLE 2. Models, parameters, and quality of the experimental semivariograms adjusted for fitting INMET air temperature. V0, nugget; V, structural variation (difference between the plateau and the nugget); (V0 + V), sill; Ro, range (km); R^2adj, model adjustment determination coefficient; r, crossed validation correlation coefficient.

Month	Models	V0	V	(V0+V)	Ro	V/(V0+V)	r	R^2	Cross validation
Jan	Exponential	0.01	1.06	1.07	149.8	99.5	0.95	0.900	y = 0,9971x + 0,0993
	spherical	0.12	0.92	1.04	128.5	88.5	0.95	0.900	y = 0,9963x + 0,1226
	Gaussian	0.26	0.81	1.07	113.9	75.7	0.95	0.895	y = 0,9945x + 0,1707
Feb	Exponential	0.01	1.13	1.14	149.6	99.6	0.95	0.897	y = 0,9945x + 0,1532
	spherical	0.12	0.98	1.10	128.2	89.1	0.95	0.898	y = 0,9942x + 0,1671
	Gaussian	0.28	0.86	1.14	113.9	75.4	0.94	0.893	y = 0,9925x + 0,2139
Mar	Exponential	0.00	1.15	1.15	149.1	99.8	0.95	0.910	y = 0,9915x + 0,2186
	spherical	0.12	1.00	1.12	132.1	89.3	0.95	0.910	y = 0,9919x + 0,2155
	Gaussian	0.28	0.87	1.15	113.9	75.7	0.95	0.906	y = 0,9907x + 0,253
Apr	Exponential	0.00	1.27	1.27	148.6	100.0	0.96	0.924	0,9901x + 0,2375
	spherical	0.13	1.09	1.22	127.7	89.3	0.96	0.925	y = 0,9904x + 0,2375
	Gaussian	0.29	0.95	1.24	109.0	76.6	0.96	0.921	y = 0,9891x + 0,2751
May	Exponential	0.00	1.31	1.31	140.6	100.0	0.97	0.936	y = 0,9908x + 0,2036
	spherical	0.14	1.16	1.30	126.9	89.2	0.97	0.937	y = 0,9906x + 0,2137
	Gaussian	0.30	1.01	1.31	108.3	77.1	0.97	0.937	y = 0,9906x + 0,2137
Jun	Exponential	0.00	1.41	1.41	140.6	100.0	0.97	0.936	y = 0,992x + 0,1702
	spherical	0.14	1.26	1.40	126.4	90.0	0.97	0.933	y = 1,0217x + 0,0362
	Gaussian	0.32	1.10	1.42	107.9	77.5	0.97	0.935	y = 0,99x + 0,2219
Jul	Exponential	0.00	1.27	1.27	140.6	100.0	0.97	0.936	y = 0,9949x + 0,1155
	spherical	0.12	1.13	1.25	126.6	90.4	0.97	0.937	y = 0,9926x + 0,1605
	Gaussian	0.29	0.98	1.27	108.1	77.2	0.97	0.934	y = 0,9909x + 0,202
Aug	Exponential	0.00	1.16	1.16	141.0	100.0	0.97	0.940	y = 0,9959x + 0,1055
	spherical	0.12	1.04	1.16	126.9	89.7	0.97	0.941	y = 0,9922x + 0,1762
	Gaussian	0.26	0.91	1.17	108.1	77.8	0.97	0.938	y = 0,9902x + 0,2244
Sep	Exponential	0.00	0.98	0.98	141.5	100.0	0.97	0.946	y = 0,994x + 0,1479
	spherical	0.09	0.86	0.95	124.2	90.5	0.97	0.947	y = 0,9914x + 0,2021
	Gaussian	0.22	0.76	0.98	109.0	77.6	0.97	0.945	y = 0,9878x + 0,2821
Oct	Exponential	0.00	0.95	0.95	140.6	100.0	0.97	0.942	y = 0,993x + 0,1768
	spherical	0.09	0.84	0.93	123.9	90.3	0.97	0.943	y = 0,9907x + 0,2297
	Gaussian	0.22	0.74	0.96	108.3	77.1	0.97	0.94	y = 0,9878x + 0,3002
Nov	Exponential	0.00	1.00	1.00	149.0	99.9	0.96	0.931	y = 0,9943x + 0,1505
	spherical	0.11	0.87	0.98	132.4	88.8	0.96	0.939	y = 0,9915x + 0,2158
	Gaussian	0.24	0.76	1.00	113.5	76.0	0.96	0.927	y = 0,9886x + 0,2881
Dec	Exponential	0.00	1.09	1.09	148.5	100.0	0.95	0.911	y = 0,9964x + 0,1035
	spherical	0.12	0.95	1.07	127.7	88.8	0.95	0.911	y = 0,9929x + 0,1834
	Gaussian	0.26	0.83	1.09	109.2	76.1	0.95	0.907	y = 0,9909x + 0,2372

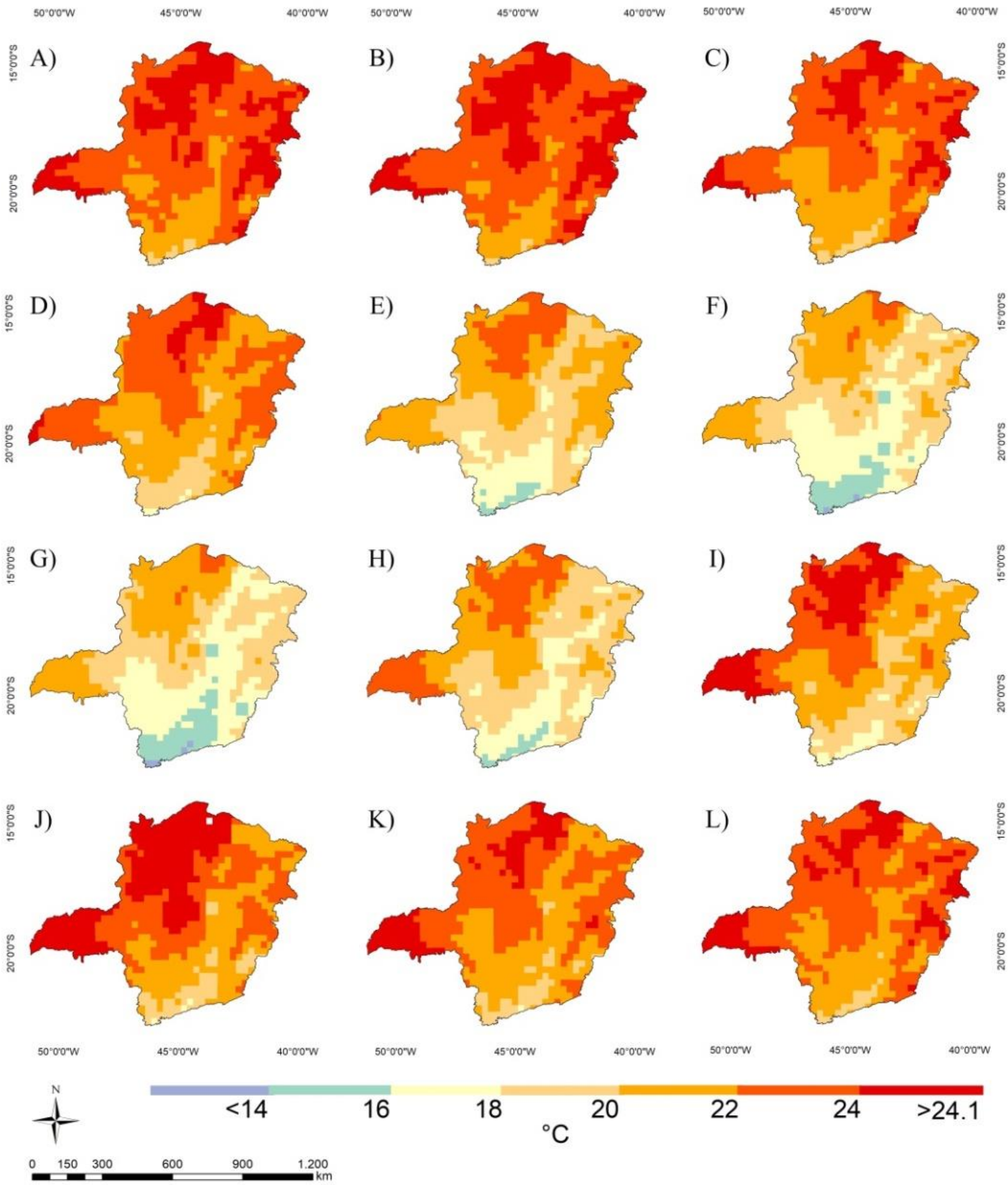

FIGURE 5. Spatial variation of air temperature from the ECMWF data for Minas Gerais, 1979-2017. A) January, B) February, C) March, D) April, E) May, F) June, G) July, H) August, I) September, J) October, K) November, and L) December. Maps are from the spherical model with one neighbour and a resolution of 0.25°.

The ECMWF system had the smallest deviations for the state, mainly in April

(Figure 6D). The deviations between the T_{INMET} and T_{ECMWF} data were largest (-2.31 °C) mainly in the Minas Gerais triangle and parts of Regiao Metropolitana, mostly in July, August, and September (Figure 6G-I). Regions such as SM and ZM had deviations <1.8 °C, but they were constant throughout the year.

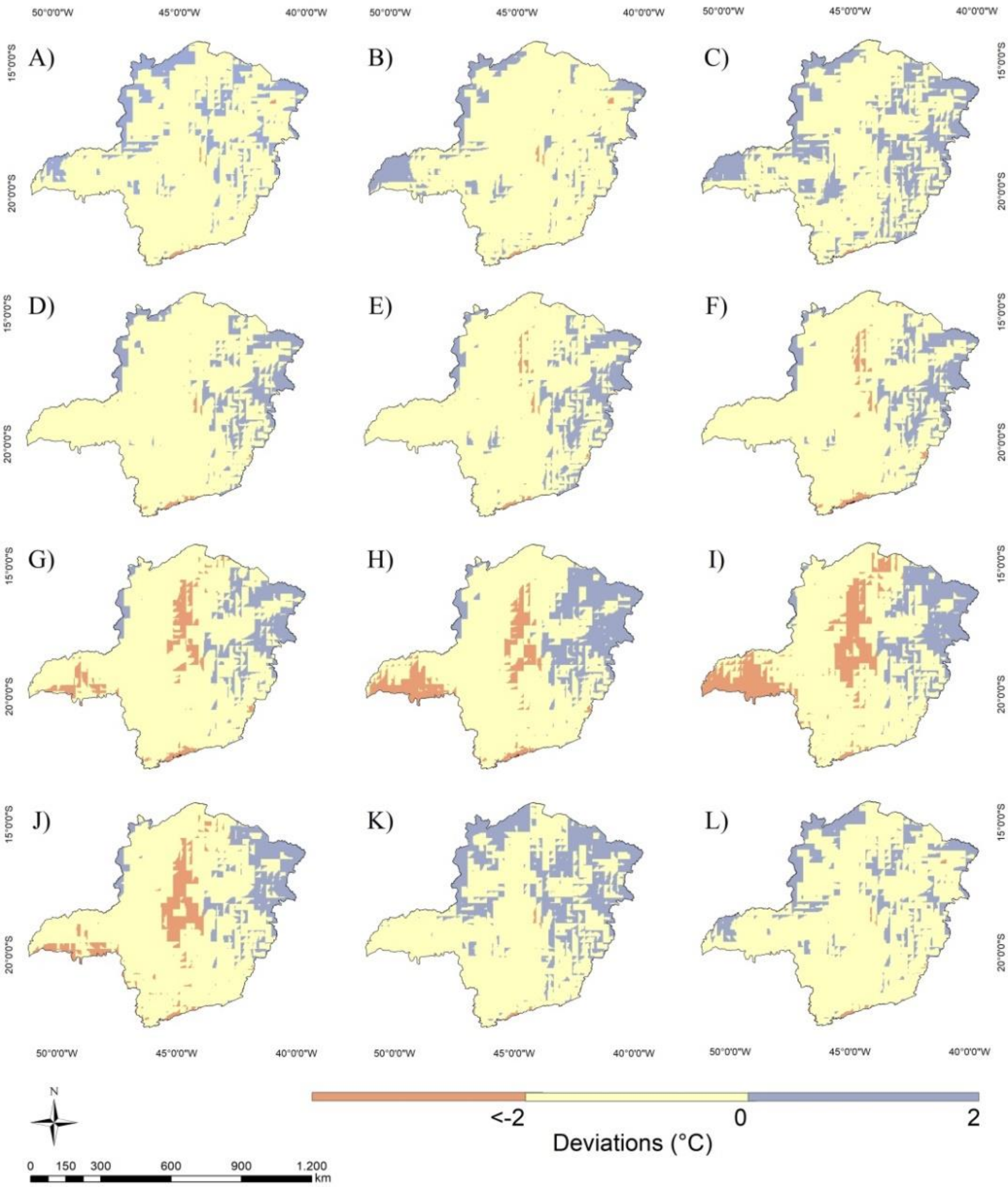

FIGURE 6. Spatialisation of the deviations between the air temperatures from the INMET surface stations and ECMWF data for Minas Gerais, 1979-2017. A) January, B) February, C) March, D) April, E) May, F) June, G) July, H) August, I) September,

J) October, K) November, and L) December. Maps are from the spherical model with one neighbour and a resolution of 0.25°.

The ECMWF system underestimated T relative to the INMET data throughout the year, especially when T was low (Figure 7). The underestimates were largest for SM, as expected, because SM is a mountainous region and had the lowest T, varying from 15 to 23 °C (Sá Júnior *et al*., 2012). For example, T for April was underestimated at values <24 °C (Figure 7D). The ECMWF system underestimated T for September in all regions except Norte de Minas (NM) (Figure 7I). The underestimates were largest for SM for each month, with high MAPEs and low R^2adj. For example, MAPE and R^2adj were 9.9% and 0.60 for January and 11.7% and 0.60 for June, respectively (Figure 8AB).

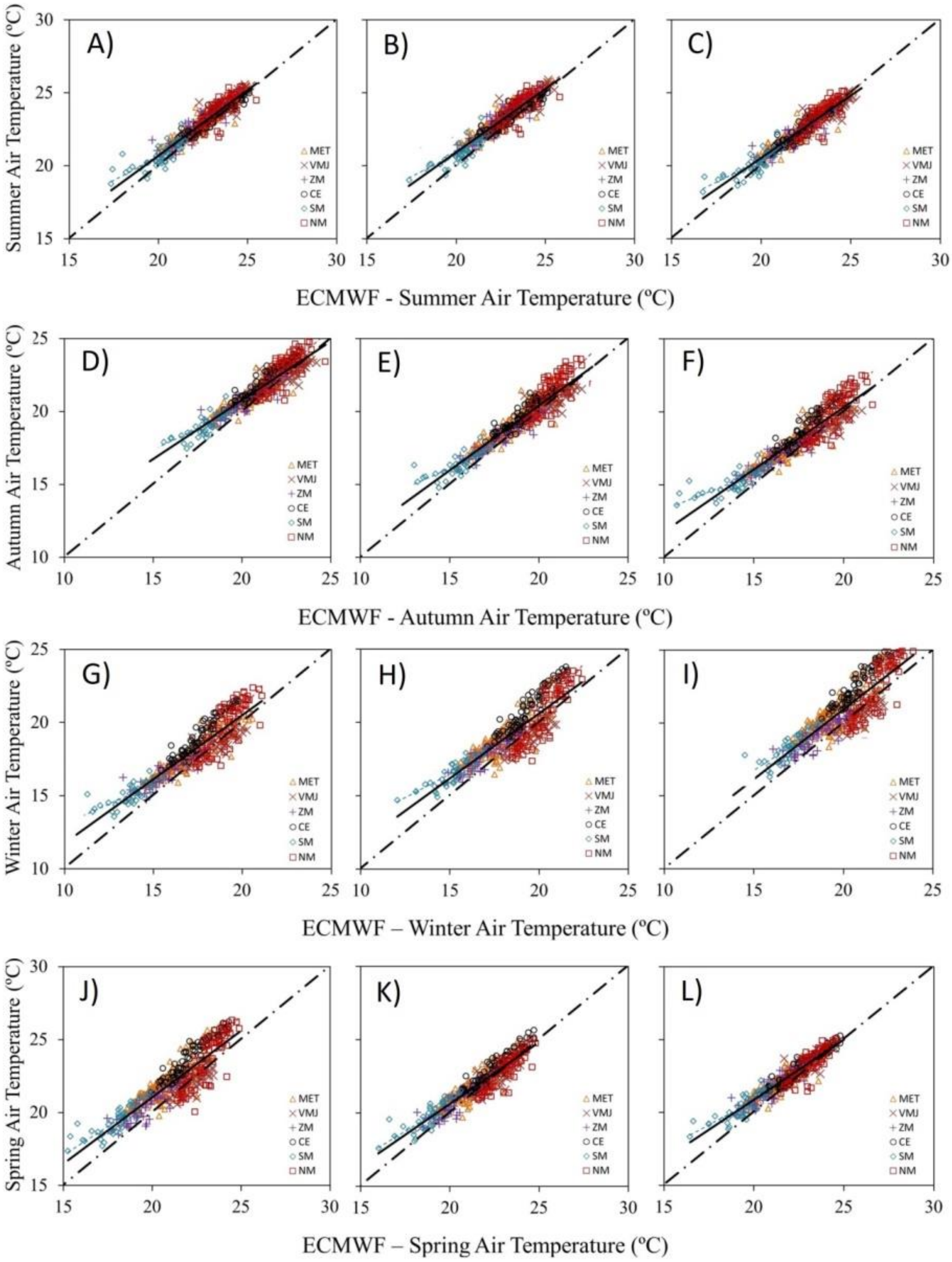

FIGURE 7. Relationship between the air temperature (°C) from the INMET surface stations and the ECMWF data for Minas Gerais, 1979-2017. MET, Regiao Metropolitana; VJM, Vale do Jequitinhonha e Mucurí; ZM, Zona da Mata; CE, Cerrado Mineiro; SM, Sul de Minas, and NM, Norte de Minas. A) January, B) February, C) March, D) April, E) May, F) June, G) July, H) August, I) September, J) October, K) November, and L) December. The dotted and dashed lines correspond to the 1:1 lines.

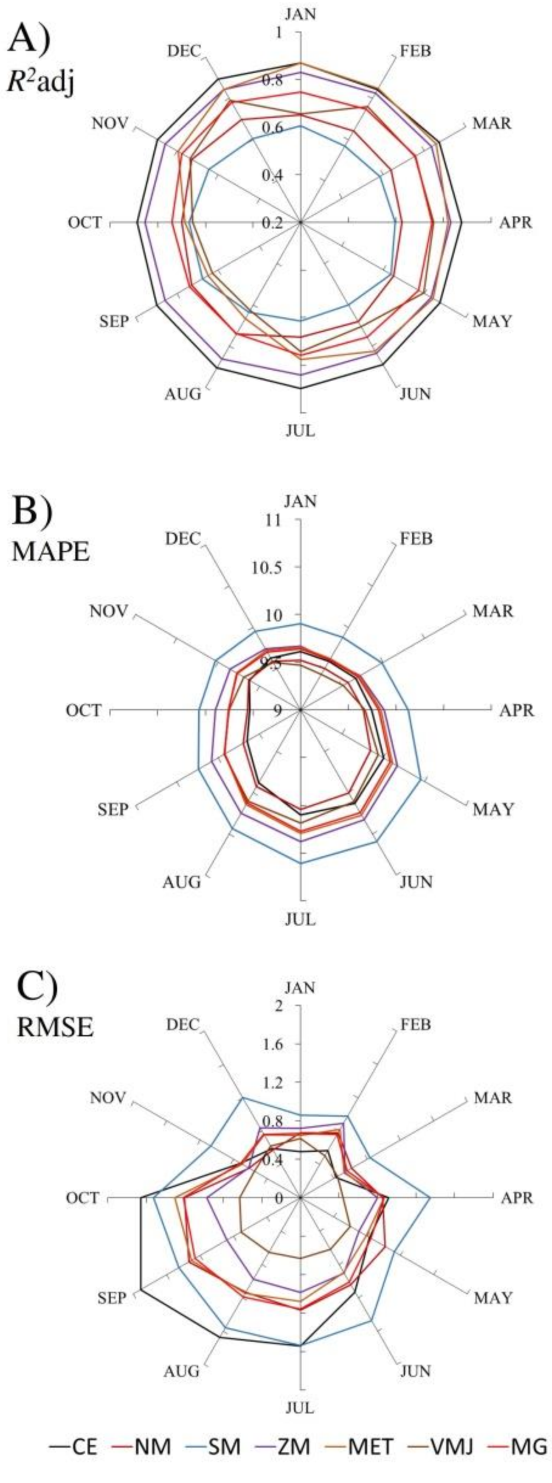

FIGURE 8. A) R^2adj (precision), B) MAPE (accuracy), and C) RMSE (accuracy) of the spatial relationship between the temperatures from the INMET surface stations and rainfall of the ECMWF system for Minas Gerais. MET, Regiao Metropolitana; VJM, Vale do Jequitinhonha e Mucurí; ZM, Zona da Mata; CE, Cerrado Mineiro; SM, Sul de Minas; NM, Norte de Minas and, MG, Minas Gerais state.

Precipitation data

ECMWF was able to monitor the temporal variability of P. The P_{ECMWF} and P_{INMET} data had similar temporal trends throughout Minas Gerais, with P highest from October to May, with a mean of 200 mm mo^{-1}, and with P lowest from April to August. Mean R^2adj, MAPE, and RMSE for the temporal variation in all periods and regions were 0.95±0.003, 14.10±17.0%, and 1.64±0.58 mm, respectively. The details of MAPE, RMSE, and R^2adj and between the temporal P_{INMET} and P_{ECMWF} data for each region are shown in Figure 9. The ECMWF system was most accurate for seasonal data in the southern region of Minas Gerais (Figure 9C), with R^2adj, MAPE, and RMSE of 0.97, 9.5%, and 0.67 mm, respectively. The MAPE of 9.5% was low, because a P of 20 mm mo^{-1} has an error of only ±1.9 mm mo^{-1}. McDonnell *et al*. (2018) found ECMWF errors <10% between P data from surface stations and ECMWF data in Ireland and ensured that ECMWF data could be used to estimate grass biomass.

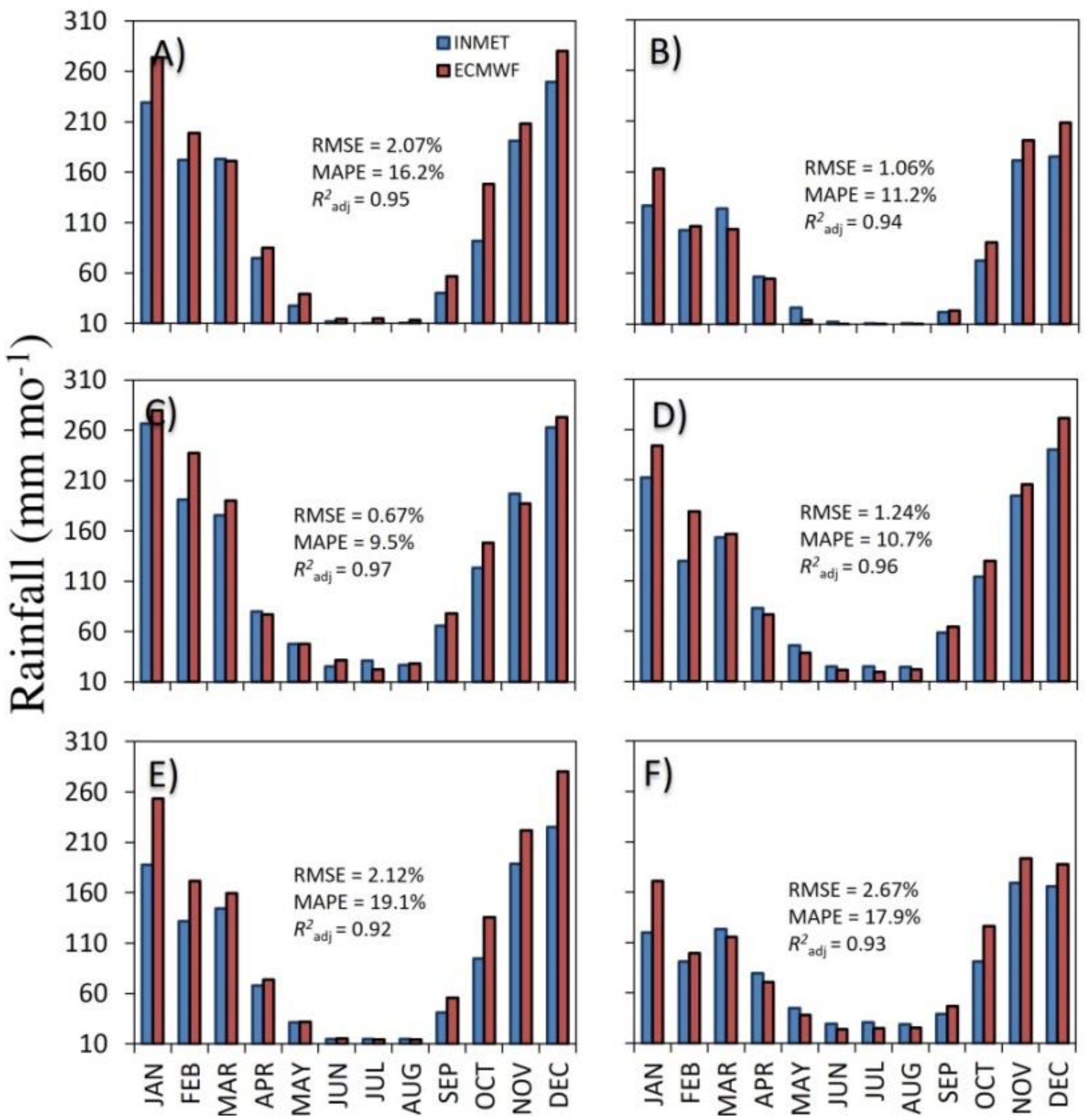

FIGURE 9. Accuracy and precision of the temporal variation of average rainfall from the INMET surface stations and ECMWF data for the main regions of Minas Gerais, 1979-2017. A) Cerrado Mineiro, B) Norte de Minas, C) Sul de Minas, D) Zona da Mata, E) Regiao Metropolitana, and F) Vale do Jequitinhonha e Mucurí.

The ECMWF system was able to represent the spatial variability of P. P was highest for December, with 315 mm in southern Minas Gerais and 175 mm in the northeast (Figure 10L). R^2adj, MAPE, and RMSE for the spatial relationship between P_{INMET} (Figure 10) and P_{ECMWF} (Figure 11) were 0.65, 10.46%, and 10.54 mm, respectively, for all sites and periods. Moraes *et al*. (2012) reported that the ECMWF system was moderately precise for P throughout most of the state of São Paulo with R^2 between 0.50 and 0.70.

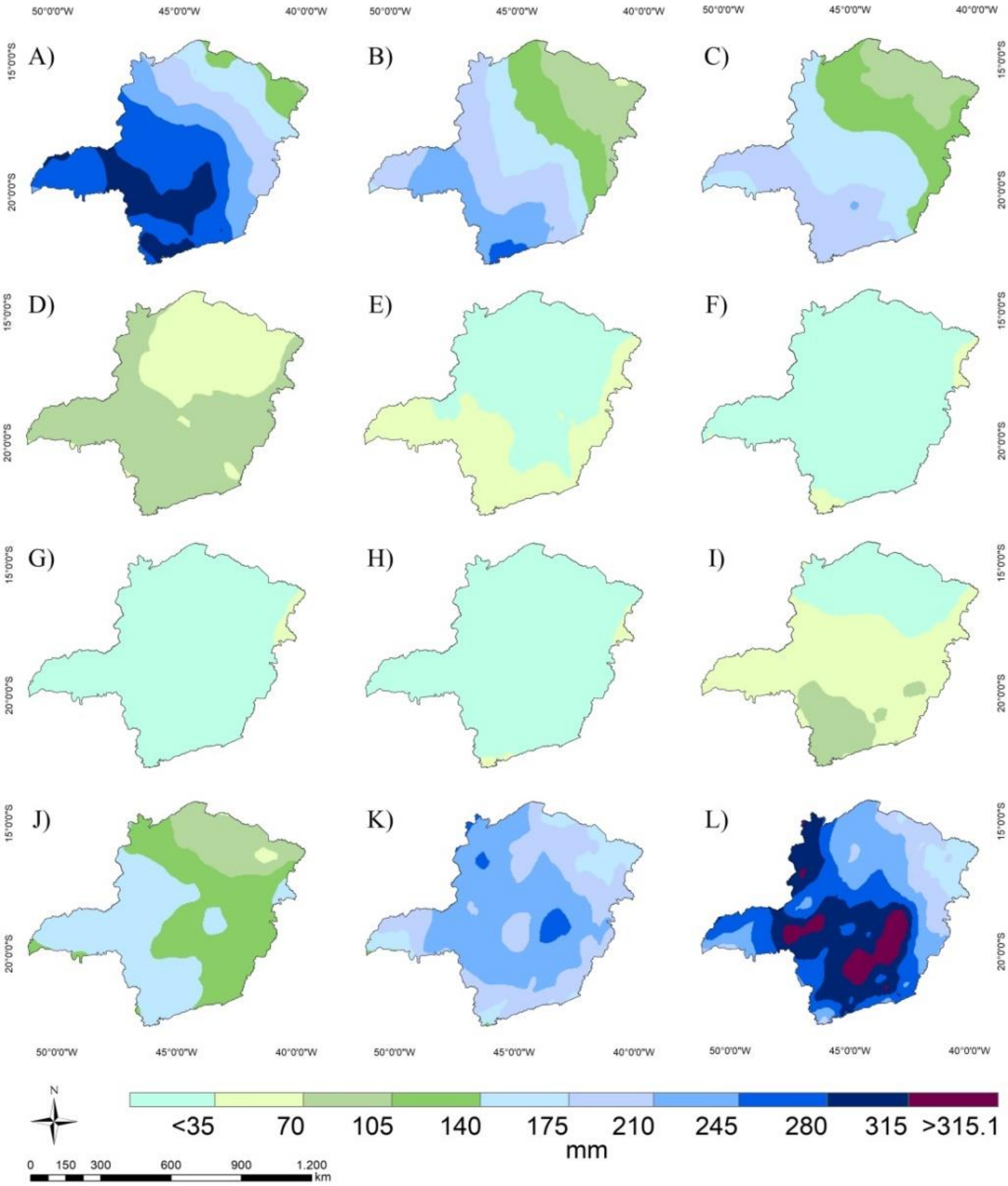

FIGURE 10. Spatial variation of the pluviometric precipitation for Minas Gerais from the INMET data, 1979-2017. A) January, B) February, C) March, D) April, E) May, F) June, G) July, H) August, I) September, J) October, K) November, and L) December. Maps are from the spherical model with one neighbour and a resolution of 0.25°.

The INMET P data in the semivariogram analysis had mean nugget, sill, and range of of 3.16±8.19 mm, 100.85±142.18 mm, and 198.92±57.74, respectively

(Table 3). The nugget adjusted in the spherical and exponential models did not vary with the monthly seasonality of P, indicating that the data did not vary. The sill varied as a function of the month. For example, P averaged 276 and 14.09 mm for the hot rainy period (November-January) and the cold dry period (May-July), respectively. P was lowest (9.95 mm) for June in the spherical model and highest (552.84 mm) for December in the Gaussian model. The range averaged 261.78 and 187.43 for the hot rainy and cold dry periods, respectively. The range was not strongly correlated with P, because P did not vary with the range.

Table 3 Models, parameters, and quality of the experimental semivariograms adjusted for fitting the INMET rainfall data. V0, nugget; V, structural variation (difference between the plateau and the nugget); (V0 + V), sill; Ro, range (km); R^2adj, model adjustment determination coefficient; r, crossed validation correlation coefficient.

Month	Models	V0	V	(V0+V)	Ro	V/(V0+V)	r	R^2	Cross validation
Jan	Exponential	0	108.6	108.68	228.2	100.0	1.00	0.996	y = 0,9961x + 0,8504
	spherical	0	115.0	115.02	193.7	100.0	1.00	0.997	y = 0,9972x + 0,6696
	Gaussian	12.42	105.5	117.94	156.2	89.5	1.00	0.997	y = 0,9968x + 0,7353
Feb	Exponential	0	103.5	103.52	320.7	100.0	1.00	0.998	y = 0,9995x + 0,3468
	spherical	0	99.32	99.32	315.5	100.0	1.00	0.998	y = 0,999x + 0,4104
	Gaussian	7.07	99.05	106.12	254.6	93.3	1.00	0.997	y = 0,9982x + 0,4826
Mar	Exponential	0	73.4	73.4	235.2	100.0	1.00	0.997	y = 1,0001x + 0,1949
	spherical	0	71.77	71.77	180.3	100.0	1.00	0.997	y = 0,9978x + 0,5493
	Gaussian	2.24	70.7	72.94	138.5	96.9	1.00	0.997	y = 0,996x + 0,7051
Apr	Exponential	0	32.83	32.83	318.7	100.0	0.99	0.985	y = 0,9945x + 0,4918
	spherical	0	34.43	34.43	226.9	100.0	0.99	0.988	y = 0,9895x + 0,8453
	Gaussian	1.57	33.95	35.52	177.1	95.6	0.99	0.987	y = 0,9848x + 1,203
May	Exponential	0	15.18	15.18	217.3	100.0	0.99	0.988	y = 1,0009x - 0,0153
	spherical	0	15.94	15.94	195.7	100.0	1.00	0.991	y = 1,0009x + 0,011
	Gaussian	0.89	15.64	16.53	154.5	94.6	1.00	0.991	y = 0,9994x + 0,0768
Jun	Exponential	0	11.54	11.54	215.6	100.0	0.99	0.984	y = 1,0143x - 0,2374
	spherical	0	9.95	9.95	124.1	100.0	0.99	0.987	y = 1,0068x - 0,0897
	Gaussian	0.71	9.02	9.73	93.1	92.7	0.99	0.989	y = 0,9982x + 0,0418
Jul	Exponential	0	17.34	17.34	259.0	100.0	0.98	0.969	y = 0,9907x + 0,1285
	spherical	0.48	15.51	15.99	165.2	97.0	0.99	0.978	y = 0,9852x + 0,2531
	Gaussian	1.8	12.86	14.66	113.9	87.7	0.98	0.969	y = 0,9855x + 0,2765
Aug	Exponential	0	15.79	15.79	196.6	100.0	0.99	0.977	y = 1,0033x - 0,0579
	spherical	0	16.57	16.57	138.2	100.0	0.99	0.985	y = 0,9989x + 0,0495
	Gaussian	0.71	17.02	17.73	112.6	96.0	0.99	0.982	y = 0,9962x + 0,1202
Sep	Exponential	0	37.74	37.74	236.8	100.0	1.00	0.916	y = 1,0004x - 0,0276
	spherical	0	39.89	39.89	207.3	100.0	1.00	0.999	y = 0,995x + 0,3417
	Gaussian	1.5	41.03	42.53	166.8	96.5	1.00	0.993	y = 0,996x + 0,2827
Oct	Exponential	0	71.37	71.37	285.0	100.0	0.99	0.984	y = 0,9958x + 0,6844
	spherical	0	72.1	72.1	215.6	100.0	0.99	0.985	y = 0,9913x + 1,3327
	Gaussian	5.84	65.37	71.21	163.8	91.8	0.99	0.985	y = 0,9859x + 2,0185
Nov	Exponential	0	176.6	176.64	245.3	100.0	0.97	0.945	y = 1,0073x - 1,6558
	spherical	10.06	169.7	179.79	221.2	94.4	0.97	0.947	y = 1,006x - 1,4311
	Gaussian	36.21	145	181.21	188.4	80.0	0.97	0.931	y = 1,0028x - 0,8807
Dec	Exponential	0	529.9	529.97	194.2	100.0	0.99	0.977	y = 0,9852x + 4,2496
	spherical	0	527.0	527.07	169.8	100.0	0.99	0.979	y = 0,9835x + 4,772
	Gaussian	32.29	520.5	552.84	135.7	94.2	0.99	0.978	y = 0,9773x + 6,3897

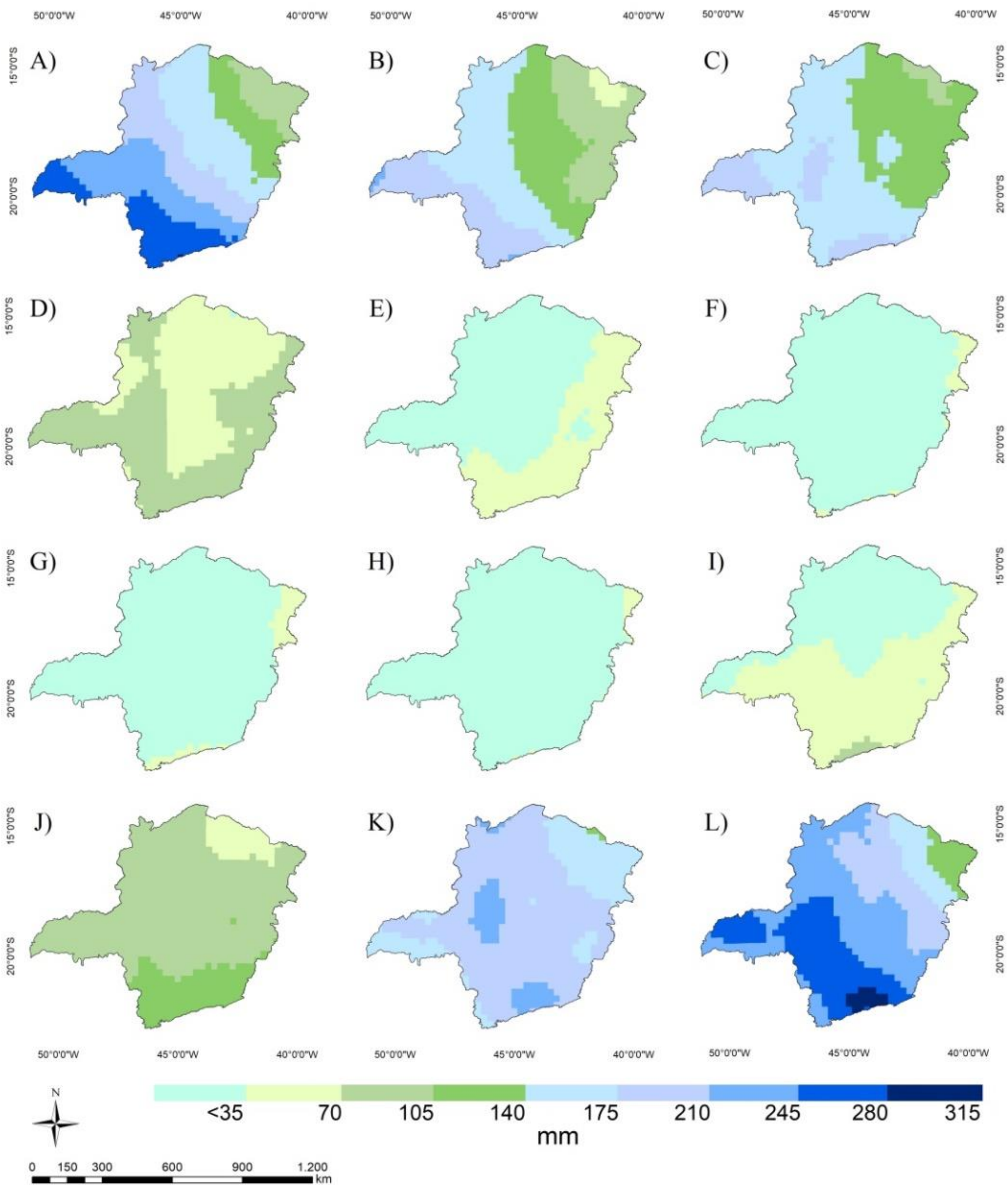

FIGURE 11. Spatial variation of precipitation from the ECMWF data for Minas Gerais, 1979-2017. A) January, B) February, C) March, D) April, E) May, F) June, G) July, H) August, I) September, J) October, K) November, and L) December. Maps are from the spherical model with one neighbour and a resolution of 0.25°.

The largest deviations between the P_{INMET} and P_{ECMWF} data averaged 75 mm mo^{-1} in spring (October-December) in the Minas Gerais triangle and central regions

(Figure 12J, L) and in summer (January and February) in the central regions (Figure 12 A, B). This result was consistent with that by Ghosh *et al.* (2018), who reported that the ECMWF system was problematic for simulating the characteristics of summer P_{ECMWF}. The deviations were lowest (5 mm mo^{-1}) for the winter (Figure 12G-I), mainly in the northeast, the Minas Triangle, and central regions.

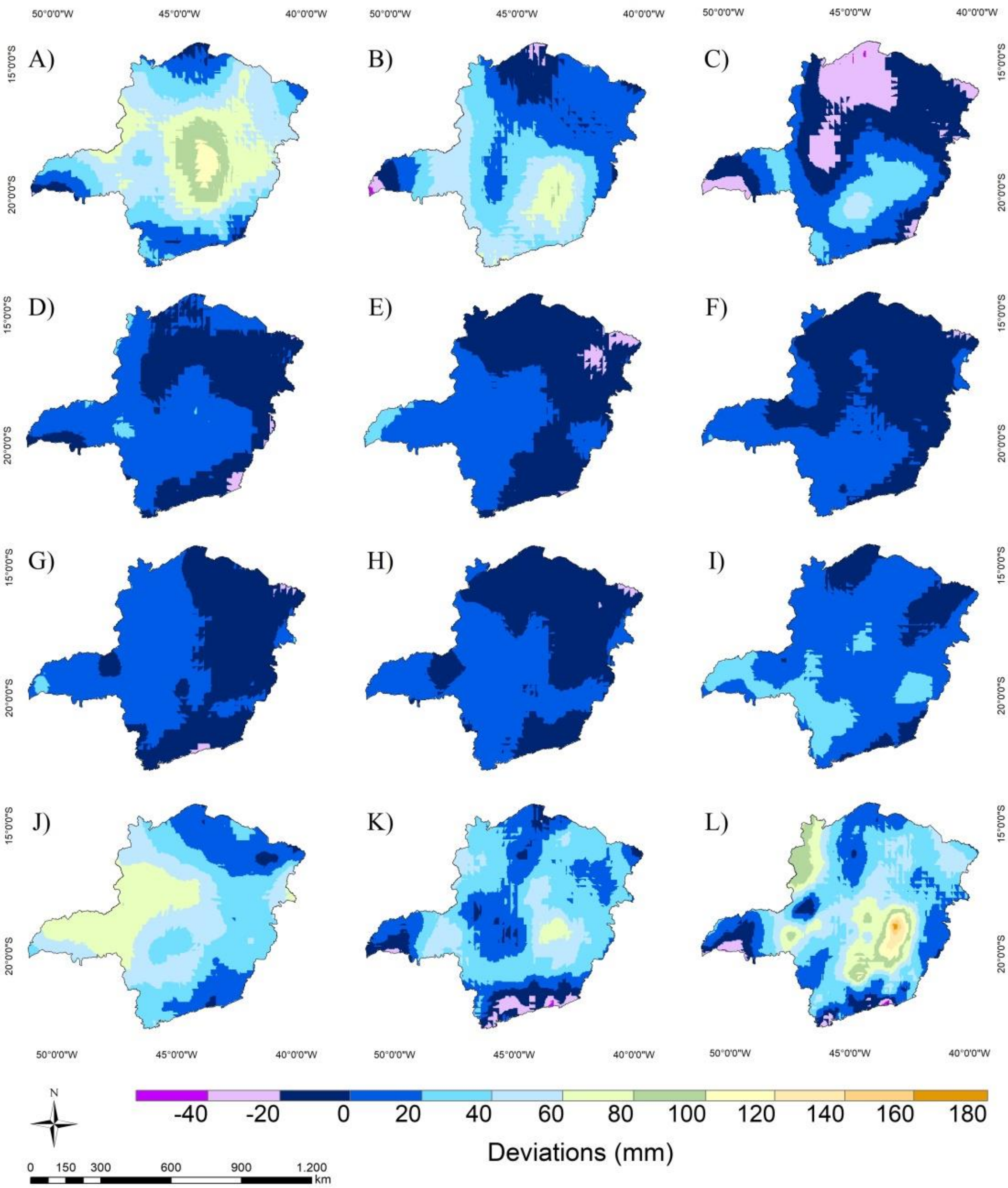

FIGURE 12. Spatialisation of deviations between precipitation data from the INMET surface stations and ECMWF data for Minas Gerais, 1979-2017. A) January, B)

February, C) March, D) April, E) May, F) June, G) July, H) August, I) September, J) October, K) November, and L) December.

The ECMWF system overestimated P for all months, especially when P increased (Figure 13). The ECMWF system underestimated the INMET values when P was low, e.g. P was underestimated for May when P was <30 mm mo^{-1}, mainly in Cerrado Mineiro (CE). P was underestimated by 60, 185, and 155 mm mo^{-1} for October, November, and December, respectively (Figure 13J-L). The accuracy was lowest when P was high, e.g. average MAPE for the summer was 20%. P was lowest for the winter, when MAPE was 9.5% (Figure 14B). McDonnell *et al*. (2018) and Ghosh et al. (2018) also reported inconsistent P_{ECMWF} results for Ireland and southern Asia, respectively.

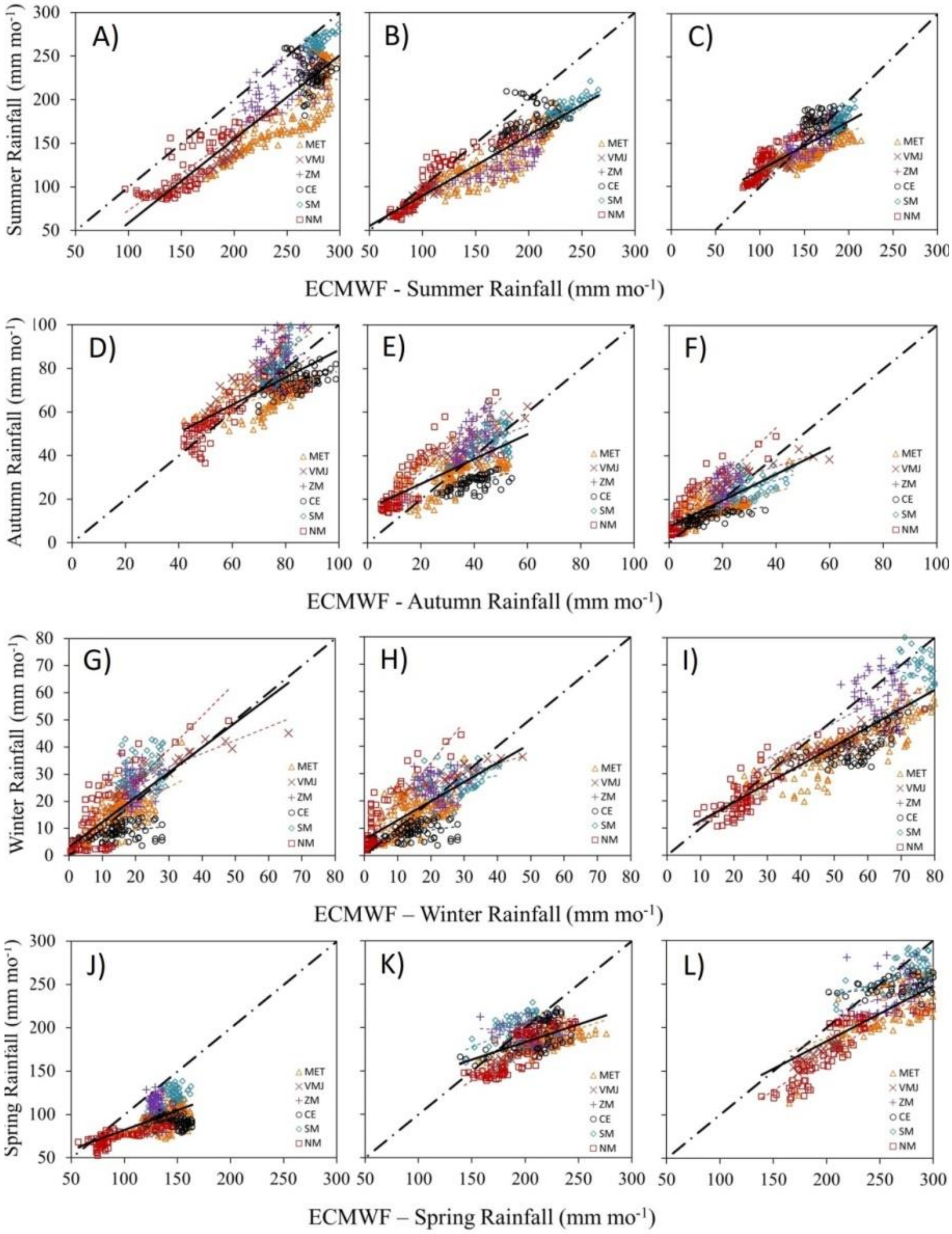

FIGURE 13. Relationship between rainfall (mm mo^{-1}) data from the INMET surface stations and ECMWF for Minas Gerais, 1979-2017. MET, Regiao Metropolitana; VJM, Vale do Jequitinhonha e Mucurí; ZM, Zona da Mata; CE, Cerrado Mineiro; SM, Sul de Minas, and NM, Norte de Minas. A) January, B) February, C) March, D) April, E) May, F) June, G) July, H) August, I) September, J) October, K) November, and L) December. The dotted and dashed lines correspond to the 1:1 lines.

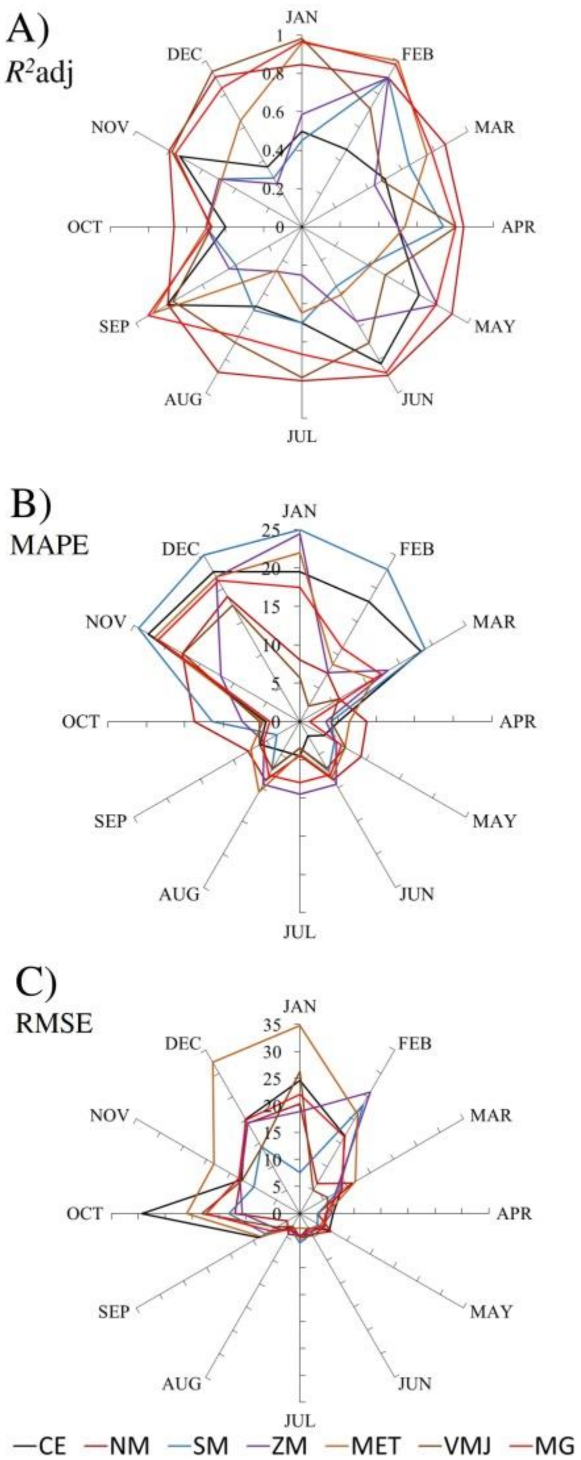

FIGURE 14. R^2adj (precision), MAPE (accuracy), and RMSE (accuracy) of the spatial relationship between the rainfall data from the INMET surface stations and the ECMWF system for Minas Gerais. MET, Regiao Metropolitana; VJM, Vale do Jequitinhonha e Mucurí; ZM, Zona da Mata; CE, Cerrado Mineiro; SM, Sul de Minas; NM, Norte de Minas, and MG, Minas Gerais state.

ECMWF data for modelling Water Balance

WB is the account of the amount of water in the soil, and soil-water storage (STO) is one of its most important components, because STO is strongly correlated with the development of agricultural crops. Conceptually, STO refers to the variation in the amount of water stored within a volume of soil over time, representing the input and outflow of water in the volume. STO decreased from March to October, with different trends and intensities for each region. For example, STO decreased from March to November in the north, by 15 mm mo^{-1} in October. STO was largest in Sul de Minas, decreasing only from May to September, mainly in august (85 mm mo^{-1}).

The comparison between STO_{INMET} and STO_{ECMWF} indicated that monthly STO was similar for the ECMWF and INMET data. STO estimates were the least accurate for VJM, with R^2adj, MAPE, and RMSE of 0.80, 21.24%, and 11.53 mm, respectively. The ECMWF data could be substituted for the INMET data for this region (Figure 15F). STO_{ECMWF} was more accurate than STO_{INMET} for southern Minas Gerais, with R^2adj, MAPE, and RMSE of 0.90, 2.35%, and 1.21 mm, respectively. A MAPE of 2.35% is considered very low, because a mean STO of 100 mm has a variation of only ±2.3 mm (Figure 15C). The deviations between STO_{INMET} and STO_{ECMWF} averaged ±10 mm mo^{-1} (Figure 16).

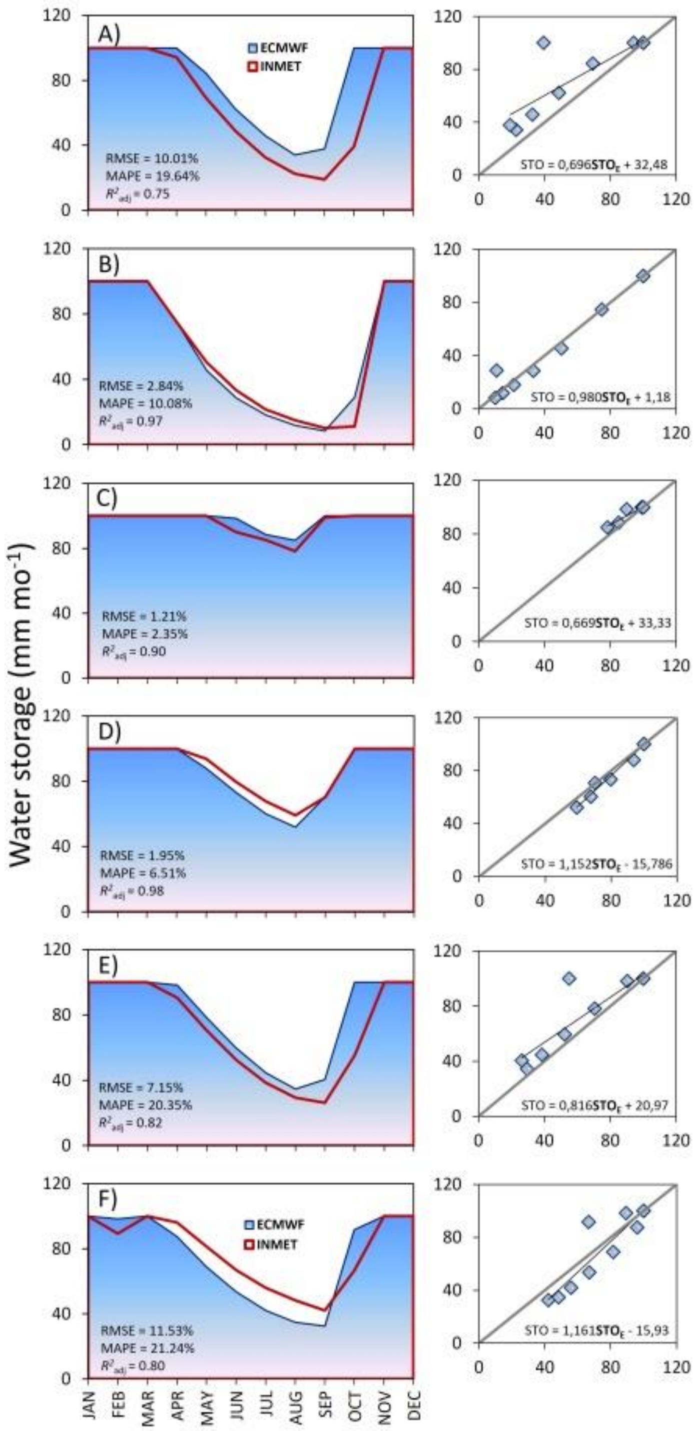

FIGURE 15. Soil-water storage from the INMET surface stations and ECMWF data for the main regions of Minas Gerais, 1979-2017. A) Cerrado Mineiro, B) Norte de Minas, C) Sul de Minas, D) Zona da Mata, E) Regiao Metropolitana, and F) Vale do Jequitinhonha e Mucurí.

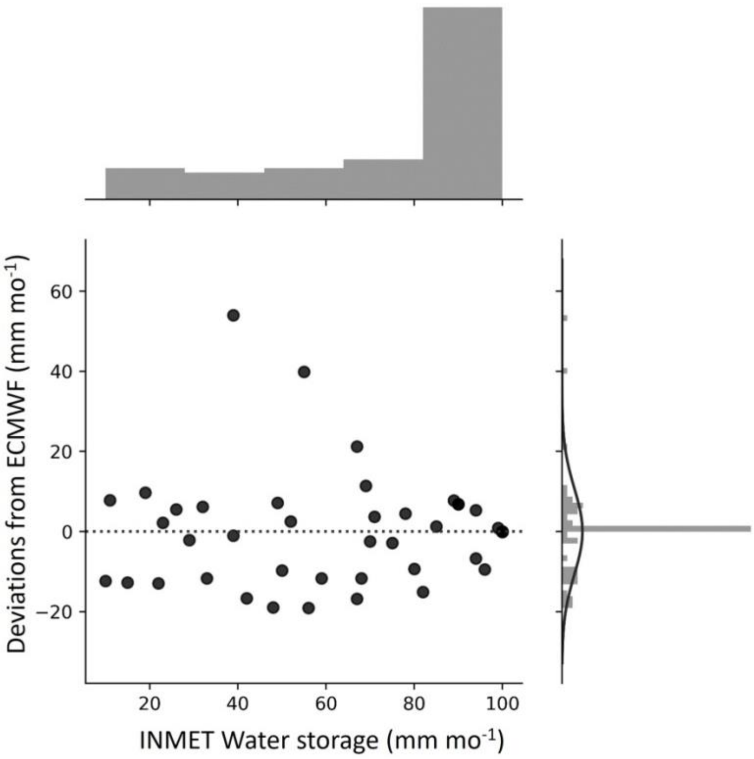

FIGURE 16. Deviations between the data for soil-water storage (mm mo^{-1}) from the INMET surface stations and ECMWF.

Water deficiency (DEF) is the WB component that influences agricultural crops the most, because DEF affects transpiration, growth, and development (Sinclair and Ludlow, 1986). DEF is defined as the difference between PET and actual evapotranspiration, i.e. the amount of water a crop loses to evapotranspiration due to a low soil STO index. DEF varied amongst the regions of the state. DEF was highest for CE and NM, with annual totals of -85.63 and -38.59 mm y^{-1}, respectively. DEF was lowest for SM and ZM, with annual totals of -2.75 and -11.67 mm y^{-1}, respectively (Figure 17).

The comparison between DEF_{INMET} and DEF_{ECMWF} indicated that DEF was similar for the ECMWF and INMET data. DEF_{ECMWF} could replace DEF_{INMET} for CE, SM, and MET but was overestimated for the other regions. The deviations between DEF_{INMET} and DEF_{ECMWF} are shown in Figure 18. The accuracy between the DEF_{ECMWF} and DEF_{INMET} data was lowest (MAPE = 19.6%) for CE (Figure 17A). A model was calibrated to adjust the ECMWF data to the INMET data ($DEF_{INMET} = 0.574 \times DEF_{ECMWF} + 0.6258$) to increase the accuracy of the data from the ECMWF.

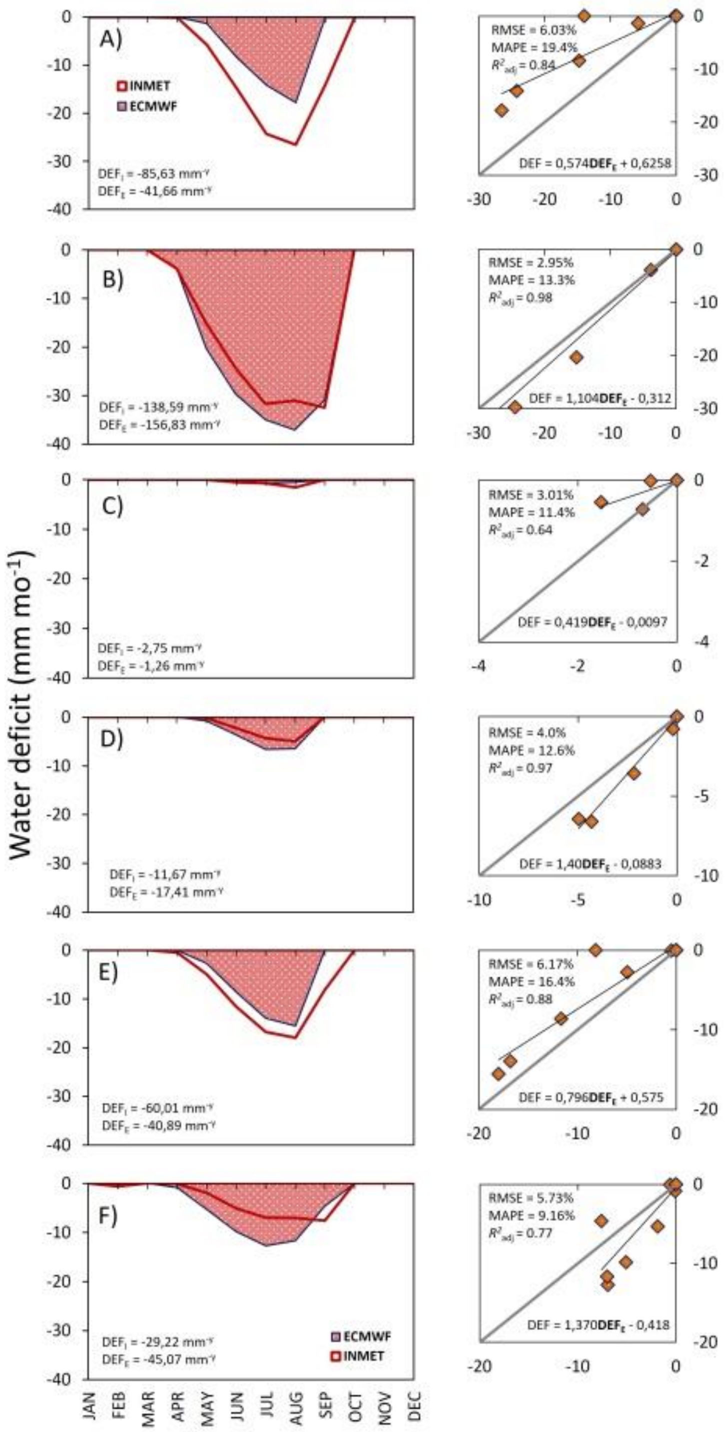

FIGURE 17. Water deficiency (mm mo^{-1}) from the INMET surface stations and ECMWF for the main regions of Minas Gerais, 1979-2017. A) Cerrado Mineiro, B) Norte de Minas, C) Sul de Minas, D) Zona da Mata, E) Regiao Metropolitana, and F) Vale do Jequitinhonha e Mucurí.

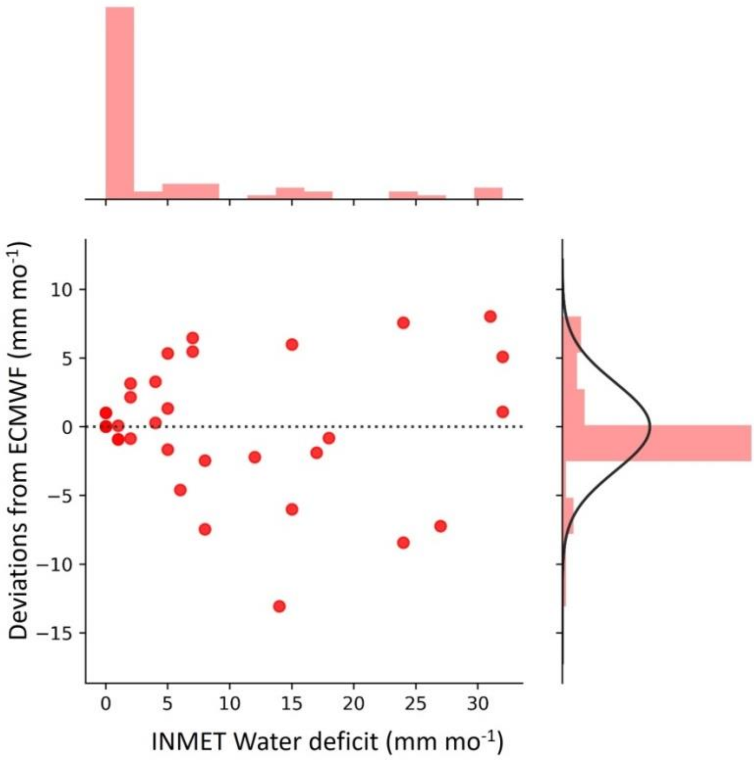

FIGURE 18. Deviations between the water-deficit data (mm mo^{-1}) from the INMET surface stations and ECMWF.

Water surplus (EXC) is another component of WB and is defined as the amount of residual water after the rainy season that is lost from a volume of soil by percolation (deep drainage) or surface runoff. EXC varied amongst the regions of the state. EXC was highest in SM from September to May, with an annual total of 793.3 mm y^{-1}. EXC was lowest in NM from October to April, with an annual total of 180.9 mm y^{-1}. The comparison between EXC_{INMET} and EXC_{ECMWF} indicated that EXC was similar for the ECMWF and INMET data. EXC_{ECMWF} was overestimated relative to the INMET data for all regions. EXC_{ECMWF} was most accurate for VJM (R^2adj = 0.92, MAPE = 8.69%, RMSE = 7.4 mm). EXC in this region occurs from October to April but mostly in November and December, with an annual total of 199.1 mm y^{-1} (Figure 19F). The deviations between EXC_{INMET} and EXC_{ECMWF} are shown in Figure 20.

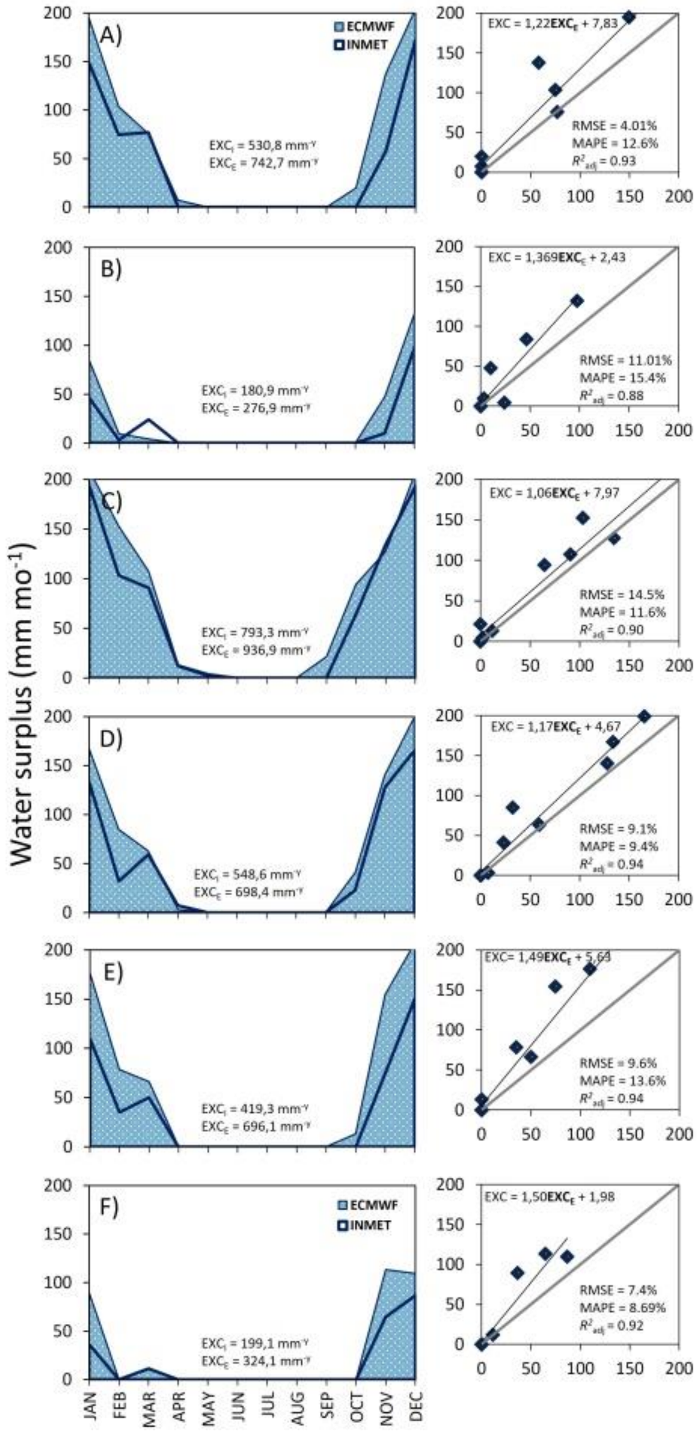

FIGURE 19. Water surplus (mm mo^{-1}) from the INMET surface stations and ECMWF for the main regions of Minas Gerais, 1979-2017. A) Cerrado Mineiro, B) Norte de Minas, C) Sul de Minas, D) Zona da Mata, E) Regiao Metropolitana, and F) Vale do Jequitinhonha e Mucurí.

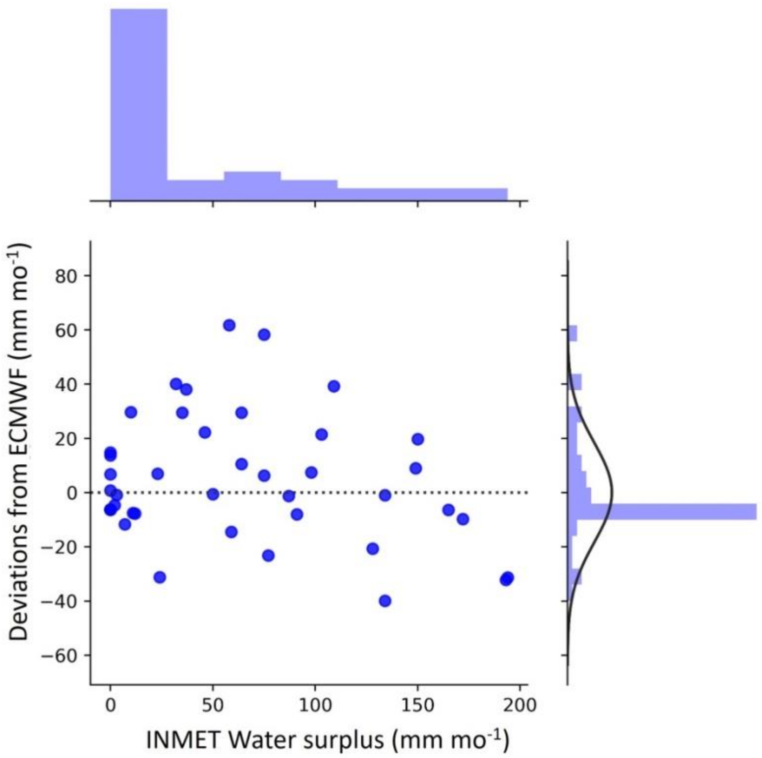

FIGURE 20. Deviations between water-surplus data (mm mo^{-1}) from the INMET surface stations and ECMWF.

Conclusions

The climatic variables from the ECMWF system were accurate and could be used for modelling climatological WB. The monthly ECMWF T and P data were more spatially and temporally accurate than the data from the INMET surface stations. The average seasonal P data were most accurate in southern Minas Gerais, with R^2adj, MAPE, and RMSE of 0.97, 9.5%, and 0.67 mm, respectively.

The ECMWF system estimated T less accurately in the coldest season (winter), with an average MAPE of 10.2%, and estimated P less accurately in the rainy season (summer), with an average MAPE of 19.9%. The ECMWF system efficiently estimated the STO, DEF, and EXC components of WB, despite the differences for T and P between seasons and regions. The mean deviations were ±10 mm mo^{-1} between STO_{INMET} and STO_{ECMWF}, ±7.6 mm mo^{-1} between DEF_{INMET}

and DEF_{ECMWF}, and ±23.6 mm mo^{-1} between EXC_{INMET} and EXC_{ECMWF}. MAPE was lowest for STO_{ECMWF} in southern Minas Gerais (1.21%), for DEF_{ECMWF} in JMV (9.16%), and for EXC_{ECMWF} also in JMV (8.69%).

Acknowledgements

This research was supported by the São Paulo Research Foundation (FAPESP, grant #: 2015/17797-4).

CAPÍTULO 5. Considerações finais

1. Atualmente existe uma preocupação mundial a respeito dos resíduos de agroquímicos nos alimentos o que tem influenciado diretamente na comercialização mundial de determinados produtos agrícolas. Com o crescente aumento dos programas de certificação no cafeeiro a preocupação com o meio ambiente tem aumentado. Acredita-se que com a utilização dos sistemas de alertas fitossanitários haverá uma redução da aplicação de agroquímicos, como inseticidas, fungicidas, e mais empresas poderão ser certificadas.

2. As modelagens de doenças e pragas em função do clima obtidas neste trabalho (Capítulos 2 e 3) podem ser utilizadas para a elaboração de sistemas de alertas fitossanitários, buscando minimizar os impactos econômicos e ambientais causados pelas incidências dessas enfermidades no cafeeiro.

3. A falta de estações meteorológicas de superfície na área em campo para aferição dos dados climáticos, não impede de fazer as modelagens tratadas neste trabalho, como observado no capítulo 4, pode ser utilizado os dados do ERA-Interim do ECMWF, uma vez que são acurados e precisos.

4. Todas as análises desta Tese serão implementadas no futuro no sistema SISMET (http://sismet.cooxupe.com.br:9000/) da Cooxupé - Cooperativa Regional de Cafeicultores em Guaxupé. Assim, pelo SISMET todos os produtores da Cooxupé terão acesso às informações elaboradas neste trabalho.

5. Com os resultados alcançados neste trabalho poderá ser utilizado como base para novas pesquisas e ser trabalhado para prestação de serviços de alertas visando uma cafeicultura economicamente e ambientalmente mais sustentável.

References

ABATZOGLOU, J. T.; DOBROWSKI, S. Z.; PARKS, S. A.; HEGEWISCH, K. C. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. Scientific data, 5, 2018. DOI: 10.1038/sdata.2017.191

ABEDETA, C.; GETU, E.; SEYOUM, E.; HINDORF, H.; BERHANE, T. Coffee leaf damaging insects occurrence in the forest coffee ecosystem of southwestern Ethiopia. African Journal of Plant Science, 9(2), 75-81, 2015.

ALBA-ALEJANDRE, I.; ALBA-TERCEDOR, J.; VEJA, F. E. 2018. Observing the devastating coffee berry borer (Hypothenemus hampei) inside the coffee berry using micro-computed tomography. Sci. Rep. 8(1): 1–9. doi: 10.1038/s41598-018-35324-4.

ALLINNE, C.; SAVARY, S.; AVELINO, J. Delicate balance between pest and disease injuries, yield performance, and other ecosystem services in the complex coffee-based systems of Costa Rica. Agriculture, Ecosystems & Environment, 222, 1-12, 2016.

ALVARES, C. A.; STAPE, J. L.; SENTELHAS, P. C.; GONÇALVES, J. L. M. Modeling monthly mean air temperature for Brazil. Theor Appl Climatol, 113, 407-427. 2013.

ANDRADE, F. T.; CASTRO JÚNIOR, L. G.; COSTA, C. H. G. Avaliação da cafeicultura pela abordagem do custeio variável em propriedades nas principais regiões produtoras do brasil. Organizações Rurais & Agroindustriais, 14(3): 356-366, 2012.

ANDROCIOLI, H. G.; HOSHINO, A. T.; MENEZES JÚNIOR, A.O.; MORAIS, H.; BIANCO, R. 2018. Coffee leaf miner incidence and its predation bay wasp in coffee intercropped with rubber trees. Coffee Sci. 13(3): 389–400.

APARECIDO, L. E. O.; ROLIM, G. S.; RICHETTI, J.; SOUZA, P. S.; JOHANN, J. A. Köppen, Thornthwaite and Camargo climate classifications for climatic zoning in the State of Paraná, Brazil. Ciência e Agrotecnologia 40(4): 405-417. 2016. doi.org/10.1590/1413-70542016404003916.

APARECIDO, L. E. O; ROLIM, G. S. Forecasting annual yield of Coffea arabica L. using monthly water deficiency. Pesquisa Agropecuária Brasileira, 53(12): 1299-1310 2018.

APARECIDO, L. E. O; ROLIM, G. S.; SOUZA, P. S. Sensitivity of newly transplanted coffee plants to climatic conditions at altitudes of Minas Gerais, Brazil. Australian Journal of Crop Science, Queensland, v.9, n.2, p.160 – 167, 2015.

APARECIDO, L.E. O.; ROLIM, G. S.; MARTORANO, L.G.; MORAES, J.R. S.C.; MENESES, K.C. Neural networks in climate spatialization and their application in the agricultural zoning of climate risk for sunflower in different sowing dates. Arch. Agron. Soil Sci. 1–16, 2019. doi: 10.1080/03650340.2019.1566715.

APARECIDO, L.E. O.; ROLIM, G. S. Forecasting of the annual yield of Arabic coffee using water deficiency. Pesq. agropec. bras., 53(12): 1299-1310, 2018.

AQUILA, V.; SWARTZ, W. H.; WAUGH, D. W.; COLARCO, P. R.; PAWSON, S.; POLVANI, L. M.; STOLARSKI, R. S. Isolating the roles of different forcing agents in global stratospheric temperature changes using model integrations with incrementally added single forcings, J. Geophys. Res. Atmos., 121, 8067–8082, 2016.

ASSIS, G. A.; GUIMARÃES, R.J.; SCALCO, M.S.; COLOMBO, A.; MORAIS, A.R.; CARVALHO, J.A.S. Correlation between coffee plant growth and yield as function of water supply regime and planting density. Biosci. J., 30(3): 666-676, 2014.

AVELINO, J.; ALLINNE, C.; CERDA, R.; WILLOCQUET, L.; SAVARY. S. Multiple-Disease System in Coffee: From Crop Loss Assessment to Sustainable Management. Annu. Rev. Phytopathol. 56(1): 611–635, 2018.

AVELINO, J.; CRISTANCHO, M.; GEORGIOU, S.; IMBACH, P.; AGUILAR, L. The coffee rust crises in Colombia and Central America (2008-2013): impacts, plausible causes and proposed solutions. 303–321, 2015.

AVELINO, J.; CRISTANCHO, M.; GEORGIOU, S.; IMBACH, P.; AGUILAR, L.; BORNEMANN, G.; LÄDERACH, P. The coffee rust crises in Colombia and Central America (2008–2013): impacts, plausible causes and proposed solutions., 7, 303–321, 2015.

BADNAKHE, M.R.; DURBHA, S.S.; JAGARLAPUDI, A.; GADE. R.M. Evaluation of Citrus Gummosis disease dynamics and predictions with weather and inversion based leaf optical model. Comput. Electron. Agric. 155, 130–141, 2018.

BAKER, K. M.; LAKE, T.; ROEHSNER, P.; SCHRANTZ, K. Forecasting Disease with 10-Year Optimized Models: Moving Toward New Digital Datasets. Agro-Geoinformatics (Agro-Geoinformatics), 2012.

BALBINO, H. T.; FORTES, L. T. G.; PARENTE, E. G. P. Avaliação do uso do modelo climático global do Centro Europeu para antecipar a estimativa do risco associado a epidemias da ferrugem Asiática da soja. Third international symposium of climatology. 2009.

BARBOSA, D.H.S.G.; VIEIRA, H.D.; SOUZA, R.M.; DIAS, P.P.; VIANA. A.P. Desenvolvimento Vegetativo e Reação de Genótipos de Coffea spp . a uma População de Meloidogyne exigua Virulenta a Cultivares Resistentes. Nematol. Bras. 31(1): 1-9, 2007.

BARBOSA, J. N.; BOREM, F. M.; ALVES, H. M. R.; VOLPATO, M. M. L.; VIEIRA, T. G. C.; SOUZA, V. C. O. Spatial distribution of coffees from minas gerais state and their relation with quality. Coffee Science, 5(3): 237-250, 2010.

BARBOSA, J. N.; BORÉM, F. M.; CIRILLO, M. Â.; MALTA, M. R.; ALVARENGA, A. A.; ALVES, H. M. R. Coffee Quality and Its Interactions with Environmental Factors in Minas Gerais, Brazil. Journal of Agricultural Science., 4(5), 2012.

BECHTOLD, P.; KOHLER, M.; JUNG, T.; DOBLAS-REYES, F.; RODWELL, M.J.; VIART, F.; BALSAMO, G. Advances in simulating atmospheric variability with the ECMWF model: From synoptic to decendal time-scales. Q.J.R. Meteorol. Soc. 134: 1337-1351, 2008.

BERGAMIN-FILHO, A.; AMORIM, L. Doenças de plantas tropicais: epidemiologia e controle econômico, 1st edn. Editora Agronômica Ceres, São Paulo. 1996.

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. Biosci. J., Uberlândia, v. 31, n. 4, p. 1004-1014, July/Aug. 2015.

BOTELHO, D.M. S.; RESENDE, M. L. V.; ANDRADE, V.T.; PEREIRA, A.A.; PATRICIO, F.R.A. Cercosporiosis resistance in coffee germplasm collection. Euphytica, 213(6), 2017. doi: 10.1007/s10681-017-1901-9.

BRAVO-MONROY, L.; POTTS, S.G.; TZANOPOULOS, J. Drivers influencing farmer decisions for adopting organic or conventional coffee management practices. Food Policy. 58:49-61, 2016.

BREIMAN, L. Random Forests. Machine Learning, 45, 5–32, 2001. doi: 10.1023/A:1010933404324.

BRUNEL-SALDIAS, N.; SEGUEL, O.; OVALLE, C.; ACEVEDO, E.; MARTÍNEZ, I. Tillage effects on the soil water balance and the use of water by oats and wheat in a Mediterranean climate. Soil and Tillage Research, 184, 68-77. 2018.

BUDYKO, M. I. The heat balance of the Earth´s surface. US Department of commerce, Washington, D.C. 1958.

BUTT, M. S.; SULTAN, M. T. Coffee and its consumption: benefits and risks. Critical Reviews in Food Science and Nutrition. 51(3), 363–373, 2011.

CAI, J., XIAO, D.; LV, L.; YE, Y. An early warning model for vegetable pests based on multidimensional data. Comput. Electron. Agric. 156: 217–226, 2019.

CAIXETA, S. L.; MARTINEZ, H. E. P.; PICANÇO, M. C.; CECON, P. R.; ESPOTI, M. D. D.; AMARAL, J. F. T. Leaf-miner attack in relation to nutrition and vigor of coffee-tree seedlings. Ciência Rural, 34(5), 1429-1435, 2004.

CAMARGO, Â. P. DE; PEREIRA, A. R. Agrometeorology of the coffee crop. World Meteorological Organization, Geneva. 1994.

CAMARGO, M. B. P. de. The impact of climatic variability and climate change on arabic coffee crop in Brazil. Bragantia, 69, 239-247, 2010.

CAMARGO, A. P.; CAMARGO, M. B. P. Definição e esquematização das fases fenológicas do cafeeiro arábica nas condições tropicais do Brasil. Bragantia. 60,65-68, 2001.

CANTELAUBE, P.; TERRES, J.M. Seasonal weather forecasts for crop yield modelling in Europe. Tellus, 57, 476–487, 2005.

CANTOR, F.; BENASSI, V. L. R. M.; FANTON, C. J. Broca-do-café, Hypothenemus hampei (Coleoptera: Scolytidae). In: VILELA, E. F.; ZUCCHI, R. A.; CANTOR, F. (Ed.). Histórico e impacto das pragas introduzidas no Brasil. Ribeirão Preto: Holos, 99-105. 2001.

CAO, F.; LIU, F.; GUO, H.; KONG, W. ZHANG, C. Fast detection of sclerotinia sclerotiorum on oilseed rape leaves using low-altitude remote sensing technology. Sensors (Switzerland) 18(12), 2018. doi: 10.3390/s18124464.

CARAMORI, P. H.; KATHOUNIAN, C. A.; MORAIS, H.; LEAL, A. C.; HUGO, R. G.; ANDROCIOLI, A. Arborização de cafezais e aspectos climatológicos. In: MATSUMOTO, S. N. (Ed.). Arborização de cafezais no Brasil. Vitória da Conquista: UESB Editions, 2004. Chapter 3, p. 19-42.

CARVALHO, A. M. D.; MENDES, A. N. G.; BOTELHO, C. E.; OLIVEIRA, A. C. B. D.; REZENDE, J. C. D.; REZENDE, R. M. Agronomic performance of coffee cultivars resistant to coffee rust in Minas Gerais State, Brazil. Bragantia, 71(4), 481-487, 2012.

CARVALHO, A. M.; CARDOSO, D. A.; CARVALHO, G. R.; CARVALHO, V. L. Comportamento de cultivares de cafeeiro sob a incidência das doenças da ferrugem e cercosporiose em dois ambientes de cultivo. Coffee Science, 12(1): 100-107, 2017.

CARVALHO, H. P.; DOURADO NETO, D.; TEODORO, R. E. F.; MELO, B. Climatological hydric balance, effective soil water storage and transpiration in coffee culture. Biosc J. 27(2), 221-229, 2011.

CARVALHO, L. G.; SEDIYAMA, G. C.; CECON, P. R.; ALVES, H. M. R. Modelo de regressão para a previsão de produtividade de cafeeiros no Estado de Minas Gerais. Revista Brasileira de Engenharia Agrícola e Ambiental, 8(2): 204-211, 2004.

CARVALHO, V. L.; CHALFOUN, S. M.; CASTRO, H. A.; CARVALHO. V. D. Influence of different yield levels on coffee rust evolution and on phenolic compounds on leaves. Ciência e Agrotecnologia, 25, 49-54, 2001.

CARVALHO, V. L.; CUNHA, R. L.; GUIMARAES, P. T. G.; CARVALHO, J. P. F. Influência do zinco na incidência de doenças do cafeeiro. Ciênc. agrotec. 32(3): 804-808, 2008.

CARVALHO-JUNIOR, W. C.; CALDERANO FILHO, B.; SILVA CHAGAS, C., BHERING, S. B.; PEREIRA, N. R.; PINHEIRO, H. S. K. Multiple linear regression and Random Forest model to estimate soil bulk density in mountainous regions. Pesquisa Agropecuária Brasileira, 51(9), 1428-1437, 2016.

CASTRO, W.; OBLITAS, J.; MAICELO, J.; AVILA-GEORGE, H. Evaluation of Expert Systems Techniques for Classifying Different Stages of Coffee Rust Infection in Hyperspectral Images. 2018.

CEGLAR, A.; TORETI, A.; LECERF, R.; VELDE, M. V. D.; DENTENER, F. Impact of meteorological drivers on regional inter-annual crop yield variability in France. Agricultural and Forest Meteorology, 216, 58–67, 2016.

CERDA, R.; AVELINO, J.; GARY, C.; TIXIER, P.; LECHEVALLIER, E. Primary and secondary yield losses caused by pests and diseases: Assessment and modeling in coffee. PLoS One 12(1): 1–17. 2017. doi: 10.1371/journal.pone.0169133.

CHALFOUN, S. M. Doenças do cafeeiro: importância, identificação e métodos de controle, First. FAEPE, Lavras. 1997.

CHAVES, E.; POZZA, E.A.; NETO, H.S.; VASCO, G.B.; DORNELAS, G.A. Temporal analysis of brown eye spot of coffee and its response to the interaction of irrigation with phosphorous levels. J. Phytopathol. 166(9): 613–622, 2018.

CONCEIÇÃO, C.H.C.; GUERREIRO-FILHO, O.; GONÇALVES, W. Fluctuation of leaf miner population in resistant arabica coffee cultivars to leaf rust. Bragantia, 64(4), 625-631, 2005.

CORNELL, J. A., BERGER, R. D. Factors that influence the coefficient of determination in single linear and nonlinear models. Phytopathology, 77, 63-70, 1987.

CORREIA, F. M.; MINGOTI, S. A.; D'ANGELO, J. V. H. Predição do número kappa de um digestor contínuo de celulose kraft usando análise de regressão múltipla. Blucher Chemical Engineering Proceedings, 1(2), 11845-11852, 2015.

COSTA, F. M. D.; ALVES, G. D. F.; SCALON, J. D.; ZACARIAS, M. S. Análise estatística das distribuições espaciais do bicho-mineiro do cafeeiro e das vespas predadoras. Coffee Science, 10(2), 149 - 157, 2015.

COUTO, F. T.; SALGADO, R.; COSTA, M. J.; PRIOR, V. Precipitation in the Madeira Island over a 10-year period and the meridional water vapour transport during the winter seasons. Int. J. Climatol. 35: 3748–3759, 2015.

CUBRY, P.; BELLIS, F.; POT, D.; MUSOLI, P.; LEROY, T. Global analysis of Coffea canephora Pierre ex Froehner from the Guineo-Congolese region reveals impacts from climatic refuges and migration effects. Genet Resour Crop Evol., 60, 483–501, 2013.

CUSTÓDIO, A. A. P.; MORAES, J. C.; CUSTÓDIO, A. A. P.; LIMA, L. A.; FARIA, M. A. The Leaf-Miner (Leucoptera Cof Eella) Incidence In Coffee Culture (Cof Ea Arabica) Under Central Pivot Irrigation. Coffee Science, 4(1):16-26, 2009.

CUSTÓDIO, A. A. P.; POZZA, E. A. P.; CUSTÓDIO, A. A. P.; SOUZA, P. E.; LIMA, L. A.; LIMA, L. M. Intensidade da ferrugem e da cercosporiose em cafeeiro quanto à face de exposição das plantas. Coffee Science, 5(3): 214-228, 2010.

DAMATTA, F. M. Ecophysiological constraints on the production of shaded and unshaded coffee: a review. Field Crops Research. 86, 99-114, 2004.

DAS, B.; NAIR, B.; REDDY, V. K.; VENKATESH. P. Evaluation of multiple linear, neural network and penalised regression models for prediction of rice yield based on weather parameters for west coast of India. Int. J. Biometeorol. 62(10): 1809–1822, 2018. doi: 10.1007/s00484-018-1583-6.

DEPPE, F.; MARTINI, L.; LONHMANN, M.; ADAMI, M. Validation studies of ECMWF precipitation data with observed SIMEPAR ground data. In: International workshop on crop monitoring and forecasting in south america, 2., Montevideo, 2006. Montevideo: South America Scientific Network on Crop Monitoring and Forecasting, 2006. p.83-92.

DONATELLI, M.; MAGAREY, R.D.; BREGAGLIO, S.; WILLOCQUET, L.; WHISH, J.P.M. Modelling the impacts of pests and diseases on agricultural systems. Agric. Syst. 155: 213–224, 2017. doi: 10.1016/j.agsy.2017.01.019.

DONG-SOON, K. I. M.; JOON-HO, L. E. E.; MYONG-SOON, Y. I. E. M. Temperature-Dependent Development of Carposina sasakii (Lepidoptera: Carposinidae) and Its Stage Emergence Models. Environmental Entomology, 30(2): 298-305, 2001.

DOORENBOS, J.; PRUITT, W.O. Guidelines for prediction crop water requirements. FAO Irrigation and Drainage Paper 24, 2º Ed, Roma 179 p., 1977.

DRIPPS, W.R.; The spatial and temporal variability of groundwater recharge within the Trout Lake basin of northern Wiscounsin. PhD Thesis, University of Wiscounsin, Madison, WI, USA. 2003.

ECHANDI, E. La chasparria de los cafetos causada por el hongo Cercospora coffeicola Berk & Cooke. Turrialba, 9(2), 54-67, 1959.

ECMWF. European Centre for Medium-Range Weather Forecasts, www.ecmwf.int. 2009.

ESQUIVEL, J. A.; PENA, F. S. Infection model for analyzing biological control of coffee rust using bacterial anti-fungal compounds. arXiv preprint arXiv:1712.08958, 2017.

FAO - Statistical Pocketbook: Coffee, Rome, 2015.

FAO, 2015. Food and Agriculture Organization of the United Nations (FAO).

FARIA, R.T.; BOWEN, W.T. Evaluation of DSSAT soil-water balance module under cropped and bare soil conditions. Braz. arch. biol. technol., 46, 489-498, 2003.

FAZUOLI, L. C.; BRAGHINI, M. T.; SILVAROLLA, M. B.; GONÇALVES, W.; MISTRO, J. C.; GALLO, P. B.; GUERREIRO FILHO, O. IAC-Obatã4739 - Arabic coffee cultivar with yellow fruits and resistant to leaf rust. Crop Breeding and Applied Biotechnology, 18(3), 330-333, 2018.

FERNANDES, F. L.; MANTOVANI, E. C.; BONFIM-NETO, H.; NUMES, V. V. Effects of Irrigation, Environmental Variability and Predatory Wasp on Leucoptera coffeella (GuérinMéneville) (Lepidoptera: Lyonetiidae), in Coffee Plants. Neotropical Entomology, 38(3), 410-417, 2009.

FERNANDES, F. L.; PICANÇO, M. C.; SILVA, R. S.; SILVA, Í. W.; FERNANDES, M. E. Controle massal da broca-do-café com armadilhas de garrafa Pet vermelha em cafeeiro. Pesquisa Agropecuária Brasileira, 49(8): 587-594, 2014.

FERNANDES, F. L.; PICANÇO, M. C.; SILVA, R. S.; SILVA, Í. W.; FERNANDES, M. E. Controle massal da broca-do-café com armadilhas de garrafa Pet vermelha em cafeeiro. Pesquisa Agropecuária Brasileira, v.49, n.8, p.587-594. 2014.

GEERAERT, L.; BERECHA, G.; HONNAY, O.; AERTS. R. Organoleptic quality of Ethiopian Arabica coffee deteriorates with increasing intensity of coffee forest management. J. Environ. Manage. 231, 282–288, 2019.

GHINI, R.; TORRE-NETO, A.; DENTZIEN, A. F.; GUERREIRO-FILHO, O.; IOST, R.; PATRÍCIO, F. R.; DAMATTA, F. M. Coffee growth, pest and yield responses to free-air CO^2 enrichment. Climatic Change, 132(2), 307-320, 2015.

GHOSH, S.; BHATLA, R.; MALL, R. K.; SRIVASTAVA, P. K.; SAHAI, A. K. Aspect of ECMWF downscaled Regional Climate Modeling in simulating Indian summer monsoon rainfall and dependencies on lateral boundary conditions. Theoretical and Applied Climatology, 1-23, 2018.

GICHIMU, B. M. Arabica coffee breeding: challenges posed by climate change. In: Proc. 10th AFCA Conference, Kampala, Uganda, 1-16, 2013.

GOURANGA, K.; ASHWANI, K. Forecasting rained rice yield with biomass of early phenophases, peak intercepted PAR and ground based remotely sensed vegetation indices. Journal of Agrometeorology. 16, 94-103, 2014.

GREE, G. Epidemiology of coffee leaf rust in the Eastern Highlands. Newsletter - Coffee Research Institute, 2, 16-20, 1993.

GUARNIERI, R.A.; PEREIRA, E.B.; MARTINS, F.R.; CHAN, C.S. Previsões de radiação solar utilizando modelo de mesoescala: Refinamento com redes neurais. 1° Congresso Brasileiro de Energia Solar, 2007.

GUERREIRO-FILHO, O. Coffee leaf miner resistance. Brazilian Journal of Plant Physiology, 18(1), 109-117, 2006.

GUJARATI, D.N., PORTER, D.C. Basics Econometrics. McGraw-Hill Education; 5 edition. 872p. 2011.

GÜMÜŞÇÜ, A.; TENEKECI, M. E.; BILGILI, A.V. Estimation of wheat planting date using machine learning algorithms based on available climate data. Sustain. Comput. Informatics Syst. 2019. doi: 10.1016/j.suscom.2019.01.010.

HADDAD, F.; MAFFIA, L. A.; MIZUBUTI, E. S. G.; TEIXEIRA. H. Biological control of coffee rust by antagonistic bacteria under field conditions in brazil. Biol Control, 49(2):114–119, 2009.

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. Agric. Food Secur. 7(1): 1–20, 2018.

HINNAH, F.D.; SENTELHAS, P.C.; MEIRA, C.A.A.; PAIVA. R.N. Weather-based coffee leaf rust apparent infection rate modeling. Int. J. Biometeorol. 62(10): 1847–1860, 2018. doi: 10.1007/s00484-018-1587-2.

HOOGENBOOM, G. Contribution of agrometeorology to the simulation of crop production and its applications. Agricultural and Forest Meteorology, v.103, n.1-2, p.137–157, 2000.

HUBER, L.; GILLESPIE, T.J. Modeling leaf wetness in relation to plant disease epidemiology. Annual Review of Phytopathology, 30, 553-577, 1992.

ICO, 2017a. International Coffee Organization (ICO) Total Production by All Exporting Countries, 31, 2017. http://www.ico.org/trade_statistics.asp?section=Statistics.

ICO, 2017b. International Coffee Organization (ICO) Country Data on the Global Coffee Trade, 2017. http://www.ico.org/profiles_e.asp.

JACKSON, D.; SKILLMAN, J.; VANDERMEER, J. Indirect biological control of the coffee leaf rust, Hemileia vastatrix, by the entomogenous fungus Lecanicillium lecanii in a complex coffee agroecosystem. Biol Control, 61(1):89–97, 2012.

JAME, Y. W.; CUTFORTH, W. Crop growth models for decision support systems. Canadian Journal of Plant Science., 76, 9-19, 1996.

JI, W.; ADAMCHUK, V.I.; CHEN, S.; MAT-SU, A.S.; ISMAIL, A. Simultaneous measurement of multiple soil properties through proximal sensor data fusion: A case study. Geoderma, 341: 111–128, 2019. doi: 10.1016/j.geoderma.2019.01.006.

JONES, J. W.; BOOTE, K. J.; JAGTAP, S. S.; WILKERSON, G. G; HOOGENBOOM, G. MISHOE, J. W. SOYGRO v.5.4- Technical documentation. Agric. Eng. Dep. Res. Rep., University of Florida, Gainesville, 1987.

JUNG, T.; MILLER, M.J.; PALMER, T.N.; TOWERS, P.; WEDI, N.; ACHUTHAVARIER, D.; ADAMS, J.M. High-resolution global climate simulations with the ECMWF model in Project Athena: Experimental design, model climate and seasonal forecast skill. Journal of Climate. 25: 3155-3172, 2012.

KIM, G.B.; JUNG, K.H.; LEE, Y.; KIM, H.J.; KIM, N. Comparison of Shallow and Deep Learning Methods on Classifying the Regional Pattern of Diffuse Lung Disease. J. Digit. Imaging, 31(4): 415–424, 2018. doi: 10.1007/s10278-017-0028-9.

KIM, K. S.; GLEASON, M. L.; AND TAYLOR, S. E. Forecasting site-specific leaf wetness duration for input to disease-warning systems. Plant Dis. v.90, p.650-656. 2006.

KITZBERGER, C. S. G.; SCHOLZ, M. B. S.; PEREIRA, L. F. P.; VIEIRA, L. G. E.; SERA, T.; SILVA, J. B. G. D.; BENASSI, M. T. Diterpenes in green and roasted coffee of Coffea arabica cultivars growing in the same edapho-climatic conditions. Journal of Food Composition and Analysis., 30, 52–57, 2013.

KOUADIO, L.; ADAMOWSKI, J. F.; PHUONG-NGUYEN, V.; DEO, R.C.; BYRAREDDY, V. Artificial intelligence approach for the prediction of Robusta coffee yield using soil fertility properties. Comput. Electron. Agric. 155, 324–338, 2018.

KRIGE, D. A statistical approach to some basic mine valuation problems on the Witwatersrand. Journal of Chemical, Metal and Mining Society of South Africa. 52:119–139, 1951.

KRISHNA, G.; SAHOO, R.N.; SINGH, P.; BAJPAI, V.; PATRA, H. Comparison of various modelling approaches for water deficit stress monitoring in rice crop through hyperspectral remote sensing. Agric. Water Manag. 213, 231–244.

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

LANDIM, P. M. B. A. Statistical analysis of geological data. 2ª . ed. revisada e ampliada – São Paulo: Editora UNESP. 2003.

LASDON, L. S.; WAREN, A. D. GRG2 user´s guide. Depto of general Business, Shchoool of Business Administration, University of Texas, Austin, TX. 1982.

LAURENTINO, E.; COSTA, J. N. M. Descrição e caracterização biológica da broca-do-café (Hypothenemus hampei, Ferrari 1867) no Estado de Rondônia. 1. ed. Porto Velho: Embrapa Rondônia, 2004. 21 p. (Embrapa Rondônia. Documentos, 90).

LI, X.; CHEN, W.; CHENG, X.; WANG, L. A comparison of machine learning algorithms for mapping of complex surface-mined and agricultural landscapes using ZiYuan-3 stereo satellite imagery. Remote sensing, v. 8, n. 6, p. 514, 2016.

LÓPES-DUQUE, S.; FERNÁNDEZ-BORRERO, O. Epidemiologia de la mancha de hierro del cafeto (Cercospora coffeicola Berk. & Cook.). Cenicafé, 3-19, 1969.

LU, J.; EHSANI, R.; SHI, Y.; ABDULRIDHA, J.; CASTRO, A.I. Field detection of anthracnose crown rot in strawberry using spectroscopy technology. Comput. Electron. Agric. 135: 289–299, 2017. doi: 10.1016/j.compag.2017.01.017.

MA, S., PEI, Z.; HE, Y. Study on simulation of rice yield with WOFOST in Heilongjiang Province. IFIP Adv. Inf. Commun. Technol. 509: 40–51, 2019.

MACHADO, J. L.; SILVA, R. A.; SOUZA, J. C. ; FIGUEIREDO, U. J.; CARVALHO, T. A. F; MATOS, C. S. M. Pragas do cafeeiro: bioecologia e manejo integrado. Informe

Agropecuário, Belo Horizonte, 35(280), 7-13, 2014.

MADDEN, L.V.; ELLIS, M. A.; LALANCETTE, N.; HUGHES, G. Evaluation of a disease warning system for downy mildew of grapes. Plant Disease, St. Paul, v.84, p.549-554, 2000.

MALAU, S.; LUMBANRAJA, P.; PANDIANGAN, S.; TARIGAN, J.R.; TINDAON. F. Performance of Coffea arabica L. In Changing Climate of North Sumatra of Indonesia. Sci. Agric. Bohem. 49(4), 340–349, 2018. doi: 10.2478/sab-2018-0041.

MARCARI, M. A.; ROLIM, G. S.; APARECIDO, L. E. O. Agrometeorological models for forecasting yield and quality of sugarcane. Australian Journal of Crop Science. 9(11), 1049-1056, 2015.

MATIELLO, J.B., SANTINATO, R., GARCIA, A.W.R., ALMEIDA, S.R., FERNANDES, D.R. Cultura de café no Brasil manual de recomendações. Rio de Janeiro-RJ e Varginha-MG, 2010. p. 280-282.

MCDONNELL, J., LAMBKIN, K., FEALY, R., HENNESSY, D., SHALLOO, L., & BROPHY, C. Verification and bias correction of ECMWF forecasts for I rish weather stations to evaluate their potential usefulness in grass growth modelling. Meteorological Applications, *25*(2): 292-301. 2018.

MEINHART, A. D.; DA SILVEIRA, T. F. F.; SILVA, R. A.; DAMIN, F. M.; BRUNS, R. E.; GODOY, H. T. Multivariate Optimization of Chlorogenic Acid Extraction From Brazilian Coffee. Food Analytical Methods, 10(9), 2943-2951, 2017.

MEIRA, C. A. A.; RODRIGUES, L. H. A.; MORAES, S. A. Modelos de alerta para o controle da ferrugem-do-cafeeiro em lavouras com alta carga pendente. Pesq. Agropec. Bras., 44(3): 233-242, 2009.

MENDONÇA, J. M. A.; CARVALHO, G. A.; GUIMARÃES, R. J.; REIS, P. R.; ROCHA, L. C. D. Produtos naturais e sintéticos no controle de Leucoptera coffeella (Guérin-Mèneville & Perrottet, 1842) (Lepdoptera: Lyonetiidae) e seus efeitos sobre a predação por vespas. Ciência e Agrotecnologia, 30(5): 892-899, 2006.

MORAES, R. A.; ARRAES, C. L. Análise de Uma Metodologia Para Preenchimento de Valores Faltantes em Dados de Precipitação, Para o Estado do Paraná. UNOPAR Cient. Exatas Tecnol., Londrina, v. 11, n. 1, p. 25-30, 2012.

MORAES, R. A.; ROCHA, J. V.; LAMPARELLI, R. A. C. Determination of total accumulated rainfall, global radiation, evapotranspiration and degree-days originated from the ECMWF model to sugar cane crop. Eng. Agríc., 34(2): 322-331, 2014.

MORAES, R. A.; ROCHA, J. V.; ROLIM, G.; LAMPARELLI, R. A. C.; MARTINS, M. Evaluation of 10-day period precipitation, maximum and minimum air temperature data from the ECMWF model in Sao Paulo state. Brasilian Journal of Irrigation and Drainage, 3, 397-407, 2012.

MOREIRA, N. B.; LIBARDI, P. L.; SALVADOR, M. M. S.; SOUSA, H. H. F. Timespace analysis of water balance components in an Oxisol. Rev. Bras. Ciênc. Solo., 38(1): 60-72, 2014.

MORETO, V.B.; ROLIM, G. S. Estimation of annual yield and quality of ‘Valencia orange’ related to monthly water deficiencies. African Journal of Agricultural Research, 10, 543 553, 2015. DOI: 10.5897/AJAR2014.9090.

NEGM, A.; FALOCCHI, M.; BARONTINI, S.; BACCHI, B. RANZI, R. Assessment of the water balance in an Alpine climate: Setup of a micrometeorological station and preliminary results. Procedia Environmental Sciences, 19:275–284. 2013.

NEYMAN, J. Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability. Philosophical Transactions of the Royal Society, 236,333–380, 1937.

NUTTER, F. W. J. Quantifying the temporal dynamics of plant virus epidemics: a review. Crop Prot 16:603–618. 1997. https://doi.org/10. 1016/S0261-2194(97)00055-0.

OLIVEIRA, R. A.; SANTOS, R. S.; RIBEIRO, A.; ZOLNIER, S.; BARBOSA, M. H. P. Estimativa da produtividade da cana-de-açúcar para as principais regiões produtoras de Minas Gerais usando-se o método ZAEI. Rev. bras. eng. agríc. ambient., v.16, n.5, pp. 549- 557, 2012.

ORTH, R.; DUTRA, E.; TRIGO, I. F.; BALSAMO, G. Advancing land surface model development with satellite-based Earth observations. Hydrol. Earth Syst. Sci. Discuss., 2016. doi:10.5194/hess-2016-628.

PEREIRA, A. R. Water balance of a podzolic soil grown with coffee trees. Bragantia, 45(2): 239-247, 1986.

PEREIRA, A. R.; CAMARGO, A. P.; CAMARGO, M. B. P. Agrometeorologia de Cafezais no Brasil. Campinas: Instituto Agronômico. p.127, 2008.

PEREIRA, A.R.; ANGELOCCI, L.R.; SENTELHAS, P.C. Agrometeorology: fundamentals and practical applications. Guaíba: Agropecuária, 2002. 478p.

PEREIRA, C. S. et al. Controle da cercosporiose e da ferrugem do cafeeiro com extrato etanólico de própolis. Revista Ceres, Viçosa, v. 55, n. 5, p. 369-376, set./out. 2008.

PEREIRA, C. S.; GUIMARÃES, R. J.; POZZA, E. A.; SILVA, A. A. Controle da cercosporiose e da ferrugem do cafeeiro com extrato etanólico de própolis. Revista Ceres, 55(5): 369-376, 2008.

PERSON, A.; GRAZZIANI, F. User guide to ECMWF forecast products. Meteorological Bulletin M3.2, version 4, March, 2007.

PICINI, A. G.; CAMARGO, M. B. P.; ORTOLANI, A. A.; FAZUOLI, L. C.; GALLO, P. B. Desenvolvimento e teste de modelos agrometeorológicos para a estimativa de produtividade do cafeeiro. Bragantia, Campinas, v.58, n.1, p.157-170, 1999.

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

PLATA-RUEDA, A.; MARTÍNEZ, L. C.; COSTA, N.C.R.; ZANUNCIO, J.C.; SENA-FERNANDES, M.E. Chlorantraniliprole–mediated effects on survival, walking abilities, and respiration in the coffee berry borer, Hypothenemus hampei. Ecotoxicol. Environ. Saf. 172: 53–58, 2019. doi: 10.1016/j.ecoenv.2019.01.063.

POURMOHAMMADALI, B.; HOSSEINIFARD, S. J.; HASSAN SALEHI, M.; SHIRANI, H.; ESFANDIARPOUR-BOROUJENI, I. Effects of soil properties, water quality and management practices on pistachio yield in Rafsanjan region, southeast of Iran. Agric. Water Manag. 213, 894–902, 2019. doi: 10.1016/j.agwat.2018.12.005.

POZZA, E. A.; CARVALHO, L. V.; CHALFOUN, S. M. Sintomas e injúrias causadas por doenças em cafeeiro. In: GUIMARÃES, R. J.; MENDES, A. N. G.; BALIZA, D. P. (Ed.). Semiologia do cafeeiro: sintomas de desordens nutricionais, fitossanitárias e fisiológicas. Lavras: UFLA, 2010. p. 68-106.

RAVINDRANATH, A.; DEVINENI, N.; LALL, U.; LARRAURI. P.C. Season-ahead forecasting of water storage and irrigation requirements-an application to the southwest monsoon in India. Hydrol. Earth Syst. Sci. 22(10): 5125–5141, 2018.

REHMAN, T.U.; MAHMUD, M.S.; CHANG, Y.K.; JIN, J.; SHIN, J. Current and future applications of statistical machine learning algorithms for agricultural machine vision systems. Comput. Electron. Agric. 156, 585–605, 2019.

RESENDE, O.; ARCANJO, R. V.; SIQUEIRA, V. C.; RODRIGUES, S. Modelagem matemática para a secagem de clones de café (Coffea canephora Pierre) em terreiro de concreto. Acta Scientiarum - Agronomy., Maringá, v.1, n.2, p.189-196, 2009.

REYES, E.I.M.; FARIAS, E.S.; SILVA, E.M.P.; FILOMENO, C.A.; PLATA, M.A.B. Eucalyptus resinifera essential oils have fumigant and repellent action against Hypothenemus hampei. Crop Prot. 116: 49–55, 2019.

RITCHIE, J.T. Soil water balance and plant water stress. In:Tsuji, G.Y.; Hoogenboom, G. Thorton, P.K. (Eds.), Understanding Options for Agricultural Production. Kluwer Academic Publishers, Dordrecht, The Nederlands, 41-55, 1998.

RODRIGUES, D.; CURE, J. R.; GUTIERREZ, A. P.; COTES, J. M.; CANTOR, F. A coffee agroecosystem model: II. Dynamics of coffee berry borer. Ecological Modelling, 248, 203–214, 2013.

RODRIGUES, R. L.; MORETTO, A. C.; GUILHOTO, J. J. M. Productive and industry structure food in paraná, from 1980 to 1995. Revista de Economia e Agronegócio, 4, 241-266, 2015.

ROLIM, G. S.; RIBEIRO, R. V.; AZEVEDO, F. A.; CAMARGO, M. B. P; MACHADO, E. Previsão do número de frutos a partir da quantidade de estruturas reprodutivas em laranjeiras. Revista Brasileira de Fruticultura, 30(1), 48-53, 2008.

RONCHI, C. P.; ARAÚJO, F. C.; ALMEIDA, W. L.; SILVA, M. A. A.; MAGALHÃES, C. E. O.; OLIVEIRA, L. B.; DRUMOND, L. C. D. Respostas ecofisiológicas de cafeeiros submetidos ao deficit hídrico para concentração da florada no Cerrado de Minas Gerais. Pesq. agropec. Bras., 50(1): 24-32, 2015.

ROWLEY, C.; CHERRILL, A.; LEATHER, S. R.; POPE, T. W. Degree-day based phenological forecasting model of saddle gall midge (Haplodiplosis marginata) (Diptera: Cecidomyiidae) emergence. Crop Protection, 102, 154-160, 2017.

SÁ JUNIOR, A.; CARVALHO, L. G.; SILVA, F. F.; ALVES, M. C. Application of the Köppen classification for climatic zoning in the state of Minas Gerais, Brazil. Theoretical and Applied Climatology, 108, 1-7, 2012.

SABINO, P.H.S.; REIS-JÚNIOR, F.A.; CARVALHO, G.A.; MANTOVANI. J.R. Nitrogen fertilizers and occurrence of Leucoptera coffeella (Guérin-Mèneville & Perrottet) in transplanted coffee seedlings. Coffee Sci. 13(3): 410–414, 2018.

SAHOO, S., RUSSO, T. A., ELLIOTT, J., & FOSTER, I. Machine learning algorithms for modeling groundwater level changes in agricultural regions of the US. Water Resources Research, v. 53, n. 5, p. 3878-3895, 2017.

SALGADO, B. G.; MACEDO, R. L. G.; CARVALHO, V. L.; SALGADO, M.; VENTURIN, N. Progresso da ferrugem e da cercosporiose do cafeeiro consorciado com grevílea, com ingazeiro e a pleno sol em Lavras - MG. Ciênc. agrotec., 31(4), 1067-1074, 2007.

SANTANA, M. F.; ZAMBOLIM, E. M.; CAIXETA, E. T.; ZAMBOLIM, L. Population genetic structure of the coffee pathogen Hemileia vastatrix in Minas Gerais, Brazil. 43(5), 473–476, 2018.

SANTOS, L. A.; SOUZA, P. E.; POZZA, E. A.; CALDEIRA, D. M.; BOTELHO, D. M. S. Nova técnica para isolar Cercospora coffeicola Berkeley & Cooke, agente etiológico da cercosporiose do cafeeiro. Coffee Science, 9(1): 142-144, 2014.

SENTELHAS, P. C.; DALLA MARTA, A.; 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, Amsterdan, v. 48, p. 392-400, 2008.

SENTELHAS, P. C.; GILLESPIE, T. J.; BATZER, J. C.; GLEASON, M. L.; MONTEIRO, J. E.; PEZZOPANE, J. R. M.; PEDRO, M. J. Spatial variability of leaf wetness duration in different crop canopies. Int. J. Biometeorol., 49, 363–370, 2005.

SENTELHAS, P. C.; GILLESPIE, T. J.; SANTOS, E. A. Leaf wetness duration measurement: comparison of cylindrical and flat plate sensors under different field conditons. Int J Biometeorol, v.51, p.265–273, 2007.

SHEKOOFA, A.; EMAM, Y.; SHEKOUFA, N.; EBRAHIMI, M.; EBRAHIMIE, E. Determining the most important physiological and agronomic traits contributing to maize grain yield through machine learning algorithms: a new avenue in intelligent agriculture. PloS one, v. 9, n. 5, 2014.

SILVA, M.G.; POZZA, E.A.; CHAVES, E.; NETO, H.S.; VASCO, G.B. Spatio-temporal aspects of brown eye spot and nutrients in irrigated coffee. Eur. J. Plant Pathol., 153(3): 931–946, 2019. doi: 10.1007/s10658-018-01611-z.

SINCLAIR, T.R.; LUDLOW, M.M. Influence of soil water supply on the plant water balance of four tropical grain legumes. Aust. J. Plant Physiol., 13, 319-340, 1986.

SINGH, A.; GANAPATHYSUBRAMANIAN, B.; SINGH, A. K.; SARKAR, S. Machine Learning for High-Throughput Stress Phenotyping in Plants. Trends in Plant Science, v.21, n.2, 2016,

SIRINUNTA, A.; AKARAPISAN, A. Screening of Antagonistic Bacteria for Controlling Cercospora coffeicola in Arabica Coffee. Journal of Agricultural Technology, 11(5), 1209-1218, 2015.

SODOUDI, S.; NOORIAN, A.; GEB, M.; REIMER, E. Daily precipitation forecast of ECMWF verified over Iran. Theoretical and Applied Climatology, 99, 39-51, 2010.

SOUZA, J. C.; REIS, P. R.; SILVA, R. A. Cafeicultor: saiba como monitorar e controlar a broca-do-café com eficiência. Lavras: Epamig, 2013b. 03 p. (Circular Técnica, 178).

SOUZA, J. C.; SOUZA, J. C.; REIS, P. R.; SILVA, R. A.; CARVALHO, T. A. F.; PEREIRA, A. B. Controle químico da broca-do-café com cyantraniliprole. Coffee Science, 8(4): 404-410, 2013.

SOUZA, V. C. O.; CUNHA, R. L.; ANDRADE, L. N.; VOLPATO, M. M. L.; CARVALHO, V. L.; ESMIN, A. A. A. Technical knowledge extraction applied to modeling of occurrence (cercospora coffeicola berkeley & cooke) coffee in the southern region of Minas Gerais. Coffee Science, 8(1), 91-100, 2013a.

SPONGOSKI, S.; REIS, P. R.; ZACARIAS, M. S. Acarofauna da cafeicultura de cerrado em Patrocínio, Minas Gerais. Ciênc. agrotec., 29(1): 9-17, 2005 .

SYVERTSEN, J. P.; GARCIA-SANCHEZ, F. Mutiple abiotic stresses occurring with salinity stress in citrus. Env. Exp. Bot. 103:128-137, 2014.

TAIZ, L.; ZEIGER, E. Livro: Fisiologia Vegetal, 4° edição, 719p., 2009.

TALAMINI, V.; POZZA, E. A.; SOUZA, P. E.; SILVA, A. M. Progresso da ferrugem e da cercosporiose em cafeeiro (coffea arabica l.) Com diferentes épocas de início e parcelamentos da fertirrigação. Ciênc. agrotec., 27(1), 141-149, 2003.

TEIXEIRA, V. L.; DEMONER, C. A.; BARBOSA, P. S. Avaliação dos diferentes tratamentos do controle químico da ferrugem e do bicho mineiro na difusão de tecnologia através do sistema de tecnologia através do sistema de metodologia e treino e visita café. SPCB - Simpósio de Pesquisa dos Cafés do Brasil 5. 2007. Águas de Lindóia, SP - Resumos Expandidos.

THORNTHWAITE, C. W. An approach towards a rational classification of climate. Geographical Review. 38, 55-94, 1948.

THORNTHWAITE, C.W.; MATHER, J.R. The water balance. Centerton: Drexel Institute of Technology, 1955. 104 p. (Publications in Climatology, v. 8, n. 1).

VAN-DAM, J.C.; HUYGEN J.; WESSELING, J.G. FEDDES, R.A.; KABAT, P.; VANWALSUM, P.E.V.GROENENDIJK, P.; Theory of SWAP v. 2.0, report #1. Department Water Resources. Wageningen Agricultural University, 167p. 1997.

VASCONCELLOS, F. C.; CAVALCANTI, I. F.A. Uma avaliação das previsões do modelo regional Eta em alta resolução para dois casos de chuva intensa ocorridos na região da Serra do Mar. Rev. Bras. de Meteorol. v. 25, n.4, p. 501-512, 2010.

VERHAGE, F.Y.F.; ANTEN, N.P.R.; SENTELHAS, P.C. Carbon dioxide fertilization offsets negative impacts of climate change on Arabica coffee yield in Brazil. Clim. Change 144(4): 671–685, 2017. doi: 10.1007/s10584-017-2068-z.

VILLACORTA, A. Alguns fatores que afetam a população estacional de Perileucoptera coffeella Guérin-Mèneville, 1842 (Lepidoptera: Lyonetiidae) no norte do Paraná, Londrina, PR. Anais da Sociedade Entomológica do Brasil,9, 23-32, 1980.

WALKER, H.E.; LEHMAN, K. A.; WALL, M.M. SIDERHURST, M. S Analysis of volatile profiles of green Hawai'ian coffee beans damaged by the coffee berry borer (Hypothenemus hampei). J. Sci. Food Agric. 99(4): 1954–1960, 2019.

WHEELER, D.L.; SCOTT, J.; SUNG DUNG, J. K.; JOHNSON, D.A. Evidence of a trans-kingdom plant disease complex between a fungus and plant-parasitic nematodes. PLoS One 14(2), 2019. doi: 10.1371/journal.pone.0211508.

XU, X.; GAO, P.; ZHU, X.; GUO, W.; DING, J. Design of an integrated climatic assessment indicator (ICAI) for wheat production: A case study in Jiangsu Province, China. Ecol. Indic. 101, 943–953, 2019. doi: 10.1016/j.ecolind.2019.01.059.

YANG, X.; MING, L.; CHUNJIANG, Z.; ZHENG, Z.; YANLIN, H. Early warning model for cucumber downy mildew in unheated greenhouses, New Zealand Journal of Agricultural Research, 50:5, 1261-1268, 2007

ZAMBOLIM, L.; VALE, F. X. R.; PEREIRA, A. A.; CHAVES, G. M. Café (Cof ea arabica L.) controle de doenças. In: VALE, F. X. R. do; ZAMBOLIM, L. Controle de doenças de plantas: grandes culturas. Viçosa: UFV; Brasília, DF: Ministério da Agricultura e Abastecimento, 2, 83-179, 1997.