



Development of warning systems for Phoma leaf spot in coffee

Humberson Rocha Silva¹, Edson Ampélio Pozza^{2*}, Aurivan Soares de Freitas³, Marcelo Loran de Oliveira Freitas⁴, Leônidas Leoni Belan⁵, Mauro Peraro Barbosa Junior², Mário Javier Ferrua Vivanco⁶ and Helon Santos Neto²

¹BioWorld Soluções Regenerativas, Janaúba, Minas Gerais, Brazil. ²Departamento de Fitopatologia, Universidade Federal de Lavras, 37200-000, Lavras, Minas Gerais, Brazil. ³Departamento de Entomologia e Fitopatologia, Universidade Federal Rural do Rio de Janeiro, Seropédica, Rio de Janeiro, Brazil. ⁴Instituto Federal de Minas Gerais, Bambuí, Minas Gerais, Brazil. ⁵Universidade Estadual da Região Tocantina do Maranhão, Imperatriz, Maranhão, Brazil. ⁶Departamento de Estatística, Universidade Federal de Lavras, Lavras, Minas Gerais, Brazil. *Author for correspondence. E-mail: edsonpozza@gmail.com

ABSTRACT. Statistical models can help in decision-making for the control of plant diseases, leading to less use of inputs, greater economy, and less negative environmental impact. Thus, this study aimed to use environmental variables to fit multiple linear regression (MLR) models for estimating the Phoma leaf spot incidence in coffee to develop a warning system. The experiment was conducted over two years (September 2013 to August 2015) with monthly disease assessments in the *Coffea arabica* L. cultivar “Catucaí amarelo 2SL”. A regular grid of 7.65 ha with 85 points delimited the area, with the points spaced 30 x 30 m. The incidence progress curve was constructed by considering the overall mean of the 85 points in each month. Fifty-two environmental variables were generated using an automatic station installed in the crop, and these variables were used in the development of the MLR models. A total of 126 models were fit, of which four were more successful in estimating disease dynamics over time. Two of these models allowed the acquisition of estimated values for disease incidence two weeks prior to the disease assessments, with high precision and accuracy. Nowadays the disease management has been performed exclusively with the use of fixed spraying schedules of fungicides. The models obtained in our research can contribute to sustainability of coffee production, to avoid unnecessary use of fungicides and become coffee cultivation more profitable.

Keywords: *Phoma* spp.; epidemiology; multiple linear regression models; meteorological data.

Received on March 2, 2024.

Accepted on July 12, 2024.

Introduction

Coffee is one of the main commodities traded worldwide. Brazil is the largest producer and exporter of coffee beans worldwide, with the main importers being the European Union (EU) and the United States (United States Department of Agriculture [USDA], 2023). The main coffee species cultivated in Brazil is *Coffea arabica* L., with a production area of 1.48 million hectares in 2023 and more than 38.9 million 60-kg bags of green coffee processed (Companhia Nacional de Abastecimento [CONAB], 2024).

A number of factors contributed to the high coffee production in Brazil, including disease control. Chemical control is employed because the main cultivars planted in most producing regions are susceptible to diseases, which can cause losses in grain yield and quality, in addition to plant death, under favorable conditions (Pozza, 2021; Sera et al., 2022). Such diseases include the Phoma leaf spot (*Phoma* spp.), which can cause losses in production of between 15 and 43%. The symptoms observed in infected plants are leaf spots, tip dieback, necrosis of rosettes, and berry mummification (Pozza et al., 2010).

Temperatures between 15 and 20°C are favorable conditions for Phoma leaf spot progress (Pozza & Alves, 2008; Santos et al., 2014; Lorenzetti et al., 2015). Humidity is another essential variable in the germination phase, germ tube growth, and pathogen infection. Furthermore, the occurrence of free water on the leaf surface is responsible for lower transpiration of the crop and water potential close to zero, providing greater internal water availability for the pathogen, which may favor its colonization (Huber & Gillespie, 1992).

This disease does not have a pattern of occurrence throughout the coffee-growing season (Santos et al., 2014), making its management more difficult. Thus, many farmers apply fungicides for prevention, using fixed spraying schedules to prevent the disease progress. This makes control costly because unnecessary applications are often made, with negative impacts on the environment and in the farmers budget. However, plant diseases occur only under favorable environmental conditions.

In this sense, statistical models capable of handling multiple variables can promote an understanding of the effects of such variables on the disease process and thus on its rational management, using environmental information to estimate future values of incidence or severity. That would enable us to define strategic periods for fungicide application, a prediction of low disease risk may result in reduced pesticide application with positive economic and sustainable environmental effects (Rossi et al., 2010; Pozza et al., 2021).

Thus, this study aimed to use environmental variables to fit multiple linear regression models for estimating the Phoma leaf spot incidence in coffee to develop a warning system.

Material and methods

The experiment was conducted for two years in a commercial crop belonging to “Neumann Kaffee Gruppe (NKG) Brazilian Farms Ltda,” located in the Santo Antônio do Amparo municipality, Minas Gerais State, Brazil. The geographical coordinates of the reference point are latitude 20°53'23.7" South, longitude 44°52'56.9" West, and altitude of 1,145.9 m. The area used in the study was 7.65 ha, cultivated with the *Coffea arabica* L. cultivar “Catucaí amarelo 2SL”, susceptible to Phoma leaf spot and resistant to coffee leaf rust (*Hemileia vastatrix* Berkeley & Broome), established at 3.7 x 0.7 m spacing, and four years old at the beginning of the study. A regular grid of 7.65 ha delimited this area with 85 points spaced 30 x 30 m apart. These points were georeferenced with a Topcon® Hiper Lite L1 L2 GPS GNSS receiver. Each georeferenced sampling point of the grid consisted of five coffee plants, with three plants in the same row, one above the central row, and the other below the central row.

Fertilization, pest and disease management in the study area were performed homogeneously and according to crop needs (Matiello et al., 2015).

MLR models were fitted to explain the behavior of the disease incidence progress curve over time. For this purpose, 24 monthly incidence assessments were carried out, beginning in September 2013 and ending in August 2015. In each plant of the grid, the first fully expanded leaf from four branches in the upper third of the canopy was evaluated randomly, with two branches on each exposure side, totaling eight leaves per plant and 40 leaves per georeferenced point. The disease incidence data at each georeferenced point were obtained dividing the number of diseased leaves by the total number of leaves, multiplied by 100.

To record the environmental variables, a Datalogger CR1000 automatic weather station, Campbell Scientific Inc.®, was installed in the field at an altitude of 1,145.04 m. Using the information collected with this station, a total of 52 variables were generated (Table 1), which were used to calculate the MLR's.

The selection of environmental variables to compose the MLR models, as well as the lag periods, were performed based on Pearson's correlation analysis ($p \leq 0.05$). The variables were lagged relative to the dates of the disease assessment in two ways:

i) the mean/cumulative values from one week prior to disease assessments, lagged by 1, 2, 3, 5, and 7 weeks prior to the evaluations, identified by 1WL, 2WL, 3WL, 5WL, and 7WL, respectively (Figure 1). For the variables “Max wind speed,” “Wind direction,” “Tmin,” “Tmean,” “RHmin,” “RHmean,” and “RHmax,” the daily means were used. For the other variables shown in Table 1, the cumulative values in the lag period were used; and

ii) the mean/cumulative values at 7, 14, 21, 28, 42, and 56 days immediately prior to the evaluations of the disease, including the evaluation day, were identified as 7P, 14P, 21P, 28P, 42P, and 56P, respectively.

Table 1. Environmental variables created using a Campbell Scientific Inc.® CR1000 Datalogger installed in the experimental area.

Variable*	Description
Wind speed max ¹	Mean of daily maximum wind speed
Wind direction ¹	Mean wind direction in the day (azimuth)
Tmin, Tmean ¹	Mean daily temperature
RHmin, RHmean, RHmax ¹	Mean daily relative air humidity
Wind speed max accum ²	Accumulated daily max speed
T15min, T15mean ²	Number of hours in the day with temperature below 15°C
T15-20min, T15-20mean ²	Number of hours in the day with temperature greater or equal to 15 and less than 20°C
T20-25min, T20-25mean ²	Number of hours in the day with temperature greater or equal to 20 and less than 25°C
T25-30min, T25-30mean ²	Number of hours in the day with temperature greater or equal to 25 and less than 30°C
RH70min, RH70mean, RH70max ²	Number of hours in the day with RH less or equal to 70%
RH80min, RH80mean, RH80max ²	Number of hours in the day with RH greater or equal to 80%
RH90min, RH90mean, RH90max ²	Number of hours in the day with RH greater or equal to 90%
NDLWmin, NDLWmean, NDLWmax ²	N° of days with leaf wetness (wet) (no wetting on the day = 0, if there was = 1)

NDLWminPAR, NDLWmeanPAR, NDLWmaxPAR ²	Number of days with PARTIAL leaf wetness (slightly wet) (no wetting = 0, if there was = 1)
NDLWmin3, NDLWmean3, NDLWmax3 ²	Number of days with leaf wetness greater or equal to 3h on day (wet)
NDLWminPAR3, NDLWmeanPAR3, NDLWmaxPAR3 ²	Number of days with PARTIAL leaf wetness greater or equal to 3h on day (slightly wet)
NDLWmin6, NDLWmean6, NDLWmax6 ²	Number of days with leaf wetness greater or equal to 6h on day (wet)
NDLWminPAR6, NDLWmeanPAR6, NDLWmaxPAR6 ²	Number of days with PARTIAL leaf wetness greater or equal to 6h on day (slightly wet)
NDLWmin9, NDLWmean9, NDLWmax9 ²	Number of days with leaf wetness greater or equal to 9h on day (wet)
NDLWminPAR9, NDLWmeanPAR9, NDLWmaxPAR9 ²	Number of days with PARTIAL leaf wetness greater or equal to 9h on day (slightly wet)
NDR ²	Number of days with rain (no precipitation on day = 0, if there was = 1)
TRD ²	Total rain in the day, in mm
DRD ²	Duration of rain on the day, in hours. Note: Smallest unit of measure is the hour.

*min, mean, and max refer to the minimum, mean, and maximum values of each variable. ¹Variable created from the mean values considered in the lag.

²Variable created from the cumulative values considered in the lag.

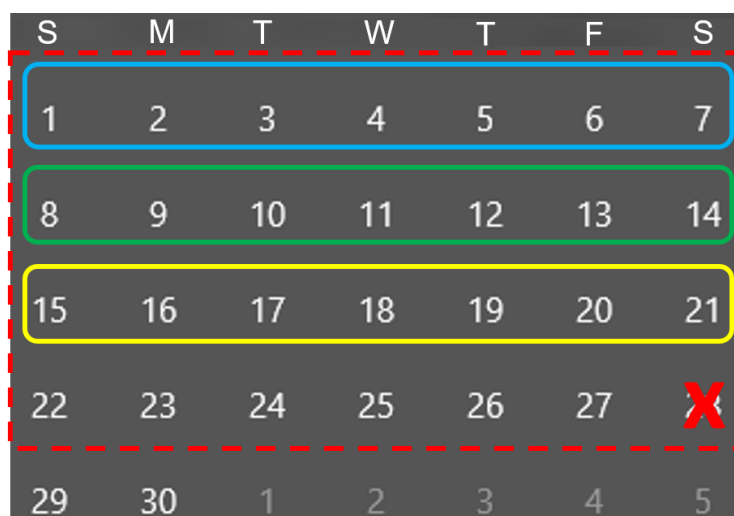


Figure 1. Representation of lags 1WL (in yellow), 2WL (in green) and 3WL (in blue), for quantification of environmental variables, considering the assessment of disease incidence on the 28th day of a hypothetical month. The red dotted line represents the days considered to obtain the variables using the 28P lag.

The selected environmental variables were incorporated into the MLR models, whose general equation is presented below (Equation 1).

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \varepsilon \quad (1)$$

where:

Y = the incidence of disease, in percentage;

x_1 , x_2 , and x_p = the environmental variables;

β_0 = the regression constant;

β_1 , β_2 ..., β_p = the partial regression parameters or coefficients; and

ε = the independent random errors.

To adjust the models, partial regression coefficients were estimated considering linear, quadratic and interaction effects between environmental variables. Initially, complete models were fitted to the data. Then, the least significant partial regression coefficients were eliminated from each model until reduced models were obtained with significant parameter estimates according to the t-test ($p \leq 0.05$), lowest Akaike Information Criterion (AIC), and highest coefficient of determination (R^2) and adjusted coefficient of determination (R^2_{Adj}).

To validate the fitted models, a progress curve was constructed with 24 monthly assessments of *Phoma* leaf spot, based on data from the plot named subgrid, composed of 121 plants located next to the CR1000 Datalogger. These plants belonged to the same cultivar and had the same age as the others, but the data collected in this area were not used to fit the regression models.

Correlation analyses were performed using Statistica 7 software (StatSoft Inc., 2004). The MLR models were fitted using the *lm* function, which is based on the ordinary least squares method, in the statistical software R version 3.4.0 (R Development Core Team, 2017).

Results

Descriptive analysis of the epidemic

Phoma leaf spot occurred sporadically over time, which is one of the basic principles for fitting predictive or decision-making support models (Campbell & Madden, 1990), ranging from 0.85% to approximately 30% of incidence in leaves (Figure 2a). The highest values of Phoma leaf spot incidence were observed in the cooler and humid periods, with incidence peaks reaching values close to 30% (Figure 2b). In 2013, the highest incidence of Phoma leaf spot was observed in October and December. In 2014, incidence peaks were observed in June, September, and November. In 2015, there was a peak in February, with a downward trend in March and April, and an increase in the months of May and June. In the second agricultural year there was greater precipitation compared to the previous year, which may have corroborated the higher incidence values (Figure 2b).

In a preliminary analysis, of the 52 environmental variables generated with the automatic station, 34 were significantly correlated with the incidence of Phoma leaf spot (data not shown). The variables related to the number of hours with temperatures between 15 and 20°C and 25 and 30°C, number of hours with relative air humidity below or equal to 70% and above or equal to 80%, and number of days with leaf wetness, were the ones most associated with the disease over time. Regarding the lag periods, in general, the 2WL, 3WL, 21P, and 28P lags returned the highest values for the correlation coefficients of the environmental variables with the disease, being, therefore, adopted for the adjustment of the MLR models.

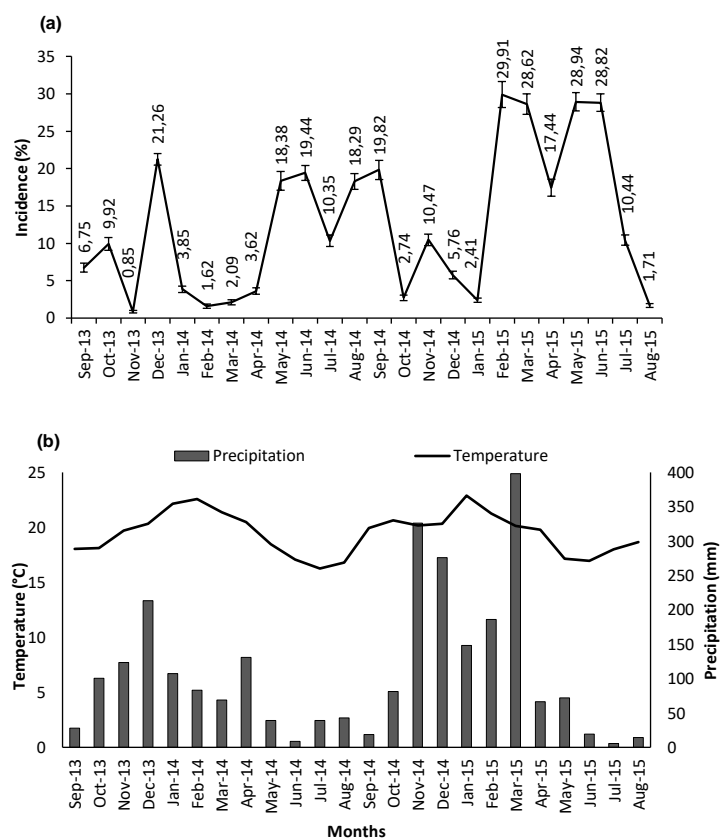


Figure 2. (a) Phoma leaf spot incidence progress curve for the overall mean of the 85 georeferenced points from September 2013 to August 2015, in a 7.65-ha area and (b) mean air temperature (°C) and cumulative precipitation (mm), 30 days prior to each monthly disease assessment. Bars in (A) indicate the standard error of the mean for each month.

In general, the variables T15-20mean, T25-30mean, RH70min, RH80max, and NDLWmax, were the ones most associated with disease progression. Thus, each of these variables was lagged by the method that provided the highest significant correlation values (“WL” lag or “P” lag). These variables were coded (x_1 , x_2 , ... x_p) for later adjustment of the MLR models (Table 2). Considering the bivariate analyses of each variable with the disease incidence, the correlation values, in module, were between 44 and 54% (Table 2). Although these values were not as high at first, when the variables were combined in the MLR models, the modeling results were promising, with high precision and accuracy.

Model fit

A total of 126 MLR models were fitted. Of this total, four had a better fit to the Phoma leaf spot progress curve. The first model presented has an intercept and 19 partial regression coefficients in a combination of linear and quadratic effects (Table 3). For this equation, the R^2 and R^2_{Adj} values were 96.82 and 81.69%, respectively. The AIC value was 66.76. In this model, four regression coefficients were not significant ($p > 0.05$). However, eliminating them caused a reduction in the R^2 and R^2_{Adj} values and an increase in the AIC value. Model 2 eliminated three nonsignificant coefficients from Model 1 and re-estimated the parameters. In this case, the R^2 and R^2_{Adj} values were 90.87 and 70.01%, respectively. The AIC increased to 86.03 (Table 3).

Model 3 was proposed, while considering the need to work only with environmental variables of mean/cumulative weekly values, with a lag of two or three weeks occurring between the evaluations and the disease. With this criterion, it was possible to forecast the disease incidence two weeks in advance. Therefore, the variables T15–20mean and T25–30mean, both with the 2WL lag, and the variables RH70min, RH80max, and NDLWmax, all with 3WL lag, were used in Model 3. Variables T25–30mean and NDLWmax, with 28P lag, were excluded from this analysis. Model 3 fit well to the monthly disease incidence dataset, with estimates of significant parameters and R^2 and R^2_{Adj} values of 96.44 and 79.51%, respectively. The AIC value for this model was 69.46 (Table 3).

Model 4 consisted solely of a manual change of the intercept of Model 3, from -3775 to -3773.5. This manual change enabled minimization of the bias errors observed in most months. No new estimates of parameters or of goodness-of-fit indicators were generated for the latter model.

Table 2. Environmental variables and lag periods, considering weeks lag (WL) and/or days immediately prior to disease assessments (P), to fit MLR models for Phoma leaf spot.

Environmental variable	Lag	Variable code in models	Correlation coefficient
T15–20mean	2WL	X ₁	+0.53*
T25–30mean	2WL	X ₂	-0.45*
T25–30mean	28P	X ₃	-0.48*
RH70min	3WL	X ₄	-0.54*
RH80max	3WL	X ₅	+0.49*
NDLWmax	3WL	X ₆	+0.44*
NDLWmax	28P	X ₇	+0.49*

*Significant by the Pearson correlation test ($p \leq 0.05$).

Table 3. Estimates of the parameters of Models 1, 2, and 3 and standard error values and significance of partial regression coefficients according to the t-test ($p \leq 0.05$).

Parameter	Model 1			Model 2			Model 3		
	Estimate	Standard error	t-test	Estimate	Standard error	t-test	Estimate	Standard error	t-test
β_0	-11790	2472	-4.771 **	-5844	1482	-3.943 **	-3775	923.3	-4.088 *
β_1	-3.707	1.144	-3.240 *	-0.967	0.4235	-2.283 *	-7.325	2.701	-2.712 *
β_2	-25.74	5.102	-5.045 **	-13.91	3.316	-4.194 **	-13.33	4.654	-2.865 *
β_3	171.9	35.67	4.818 **	85.96	21.39	4.019 **	60.76	12.81	4.743 **
β_4	171.6	36.36	4.720 **	83.8	21.54	3.890 **	51.06	11.81	4.323 *
β_5	-167.1	37.33	-4.476 *	-80.98	25.03	-3.236 *	86.75	18.26	4.750 **
β_6	0.01503	0.006322	2.378 ns	-0.3089	0.07629	-4.049 **	-0.09681	0.01691	-5.727 **
β_7	-0.6111	0.1259	-4.855 **	-0.2941	0.07688	-3.826 **	-0.2071	0.03596	-5.759 **
β_8	-0.6117	0.1312	-4.663 **	0.01706	0.004704	3.626 **	-0.2663	0.04982	-5.344 **
β_9	-0.109	0.06336	-1.720 ns	0.06739	0.02403	2.804 *	-0.2261	0.04686	-4.824 **
β_{10}	0.03128	0.006644	4.708 **	0.08336	0.02	4.168 **	-0.2914	0.05218	-5.583 **
β_{11}	0.005634	0.003975	1.417 ns	0.07385	0.01832	4.031 **	0.169	0.03228	5.237 **
β_{12}	0.04302	0.02451	1.755 ns	0.07658	0.02184	3.506 **	0.196	0.0359	5.459 **
β_{13}	0.1541	0.0305	5.053 **	-0.5886	0.1507	-3.906 **	-0.7529	0.1598	-4.711 **
β_{14}	0.1437	0.02943	4.884 **	0.459	0.1568	2.927 *	0.2597	0.04801	5.409 **
β_{15}	0.06652	0.02101	3.166 *	0.479	0.1512	3.167 *	0.2705	0.05063	5.343 **
β_{16}	-1.201	0.2537	-4.733 **	-1.883×10^{-7}	4.923×10^{-8}	-3.826 **	-0.8367	0.2189	-3.822 *
β_{17}	1.022	0.2421	4.222 *	NA	NA	NA	-0.4969	0.09618	-5.166 **
β_{18}	1.063	0.2449	4.340 *	NA	NA	NA	-0.2886	0.0832	-3.469 *
β_{19}	-2.623×10^{-7}	4.951×10^{-8}	-5.297 **	NA	NA	NA	2.345×10^{-6}	5.821×10^{-7}	4.029 *
R^2		96.82%			90.87%			96.44%	
R^2_{Adj}		81.69%			70.01%			79.51%	
AIC		66.76			86.03			69.46	

ns, *, and ** - Not significant, significant at the 5% probability level, and significant at the 1% probability level, respectively, by the t-test. NA - Not applicable, AIC - Akaike information criterion, R^2 - coefficient of determination and R^2_{Adj} - adjusted coefficient of determination.

The models are described in Equations 2, 3, and 4.

Model 1: (2)

$$\text{Incidence} = -11790 - 3.707x_1 - 25.74x_3 + 171.9x_4 + 171.6x_5 - 167.1x_7 + 0.01503x_1^2 - 0.6111x_4^2 - 0.6117x_5^2 - 0.109x_7^2 + 0.03128(x_1x_3) + 0.005634(x_1x_5) + 0.04302(x_1x_7) + 0.1541(x_3x_4) + 0.1437(x_3x_5) + 0.06652(x_3x_7) - 1.201(x_4x_5) + 1.022(x_4x_7) + 1.063(x_5x_7) - 0.0000002623(x_1x_3x_4x_5x_7)$$

Model 2: (3)

$$\text{Incidence} = -5844 - 0.967x_1 - 13.91x_3 + 85.96x_4 + 83.8x_5 - 80.98x_7 - 0.3089x_4^2 - 0.2941x_5^2 + 0.01706(x_1x_3) + 0.06739(x_1x_7) + 0.08336(x_3x_4) + 0.07385(x_3x_5) + 0.0765(x_3x_7) - 0.5886(x_4x_5) + 0.459(x_4x_7) + 0.479(x_5x_7) - 0.0000001883(x_1x_3x_4x_5x_7)$$

Model 3: (4)

$$\text{Incidence} = -3775 - 7.325x_1 - 13.33x_2 + 60.76x_4 + 51.06x_5 + 86.75x_6 - 0.09681x_1^2 - 0.2071x_2^2 - 0.2663x_4^2 - 0.2261x_5^2 - 0.2914(x_1x_2) + 0.169(x_1x_4) + 0.196(x_1x_5) - 0.7529(x_1x_6) + 0.2597(x_2x_4) + 0.2705(x_2x_5) - 0.8367(x_2x_6) - 0.4969(x_4x_5) - 0.2886(x_4x_6) + 0.000002345(x_1x_2x_4x_5x_6)$$

The incidence values in each month were estimated by replacing the variables x_1, x_2, \dots, x_7 in the models with the values recorded in the automatic station.

Model 3 showed good precision in estimating incidence values over the two years of assessments, and in terms of accuracy, it showed a small underestimation bias most of the time (Figure 3a). With the modification of the intercept value of this model, Model 4 was originated, which was, at the same time, highly precise and accurate (Figure 3b).

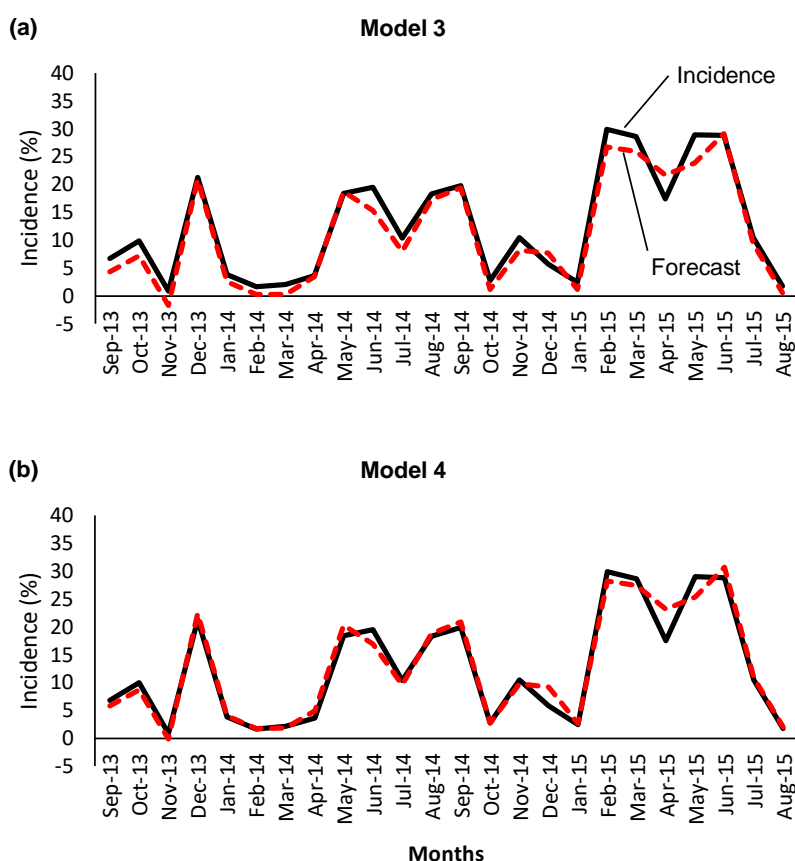


Figure 3. Phoma leaf spot progress curve in the coffee crop (mean of the 85 georeferenced points) and estimates obtained with (a) Model 3 and (b) Model 4.

Validation was performed only for models 3 and 4 because they allow estimating the disease two weeks in advance, according to the criteria adopted for the construction of these models (Figure 4). All disease peaks fitted to models 3 and 4 were observed in the progress curve constructed with the subgrid data (Figure 4). However, the incidence values of this plot were higher than the model estimates, especially in the months of December 2013 and 2014 and in February, March, and June 2015, when the highest peaks and, consequently, the largest deviations were recorded.

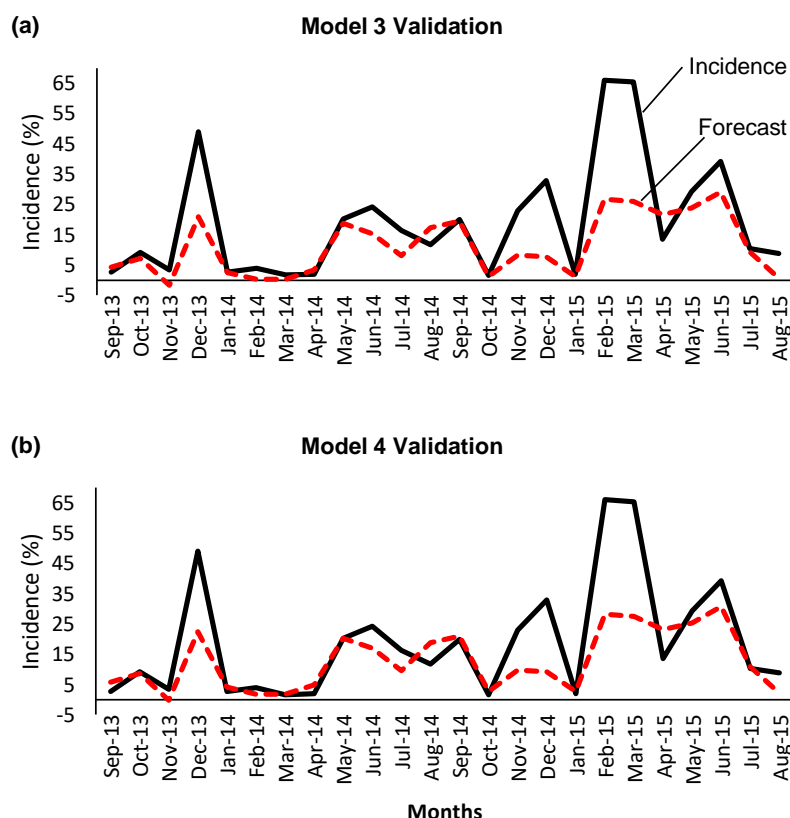


Figure 4. Phoma leaf spot progress curve in the subgrid (plot composed of 121 plants) and estimates obtained with validation of (a) Model 3 and (b) Model 4.

Discussion

Biological interpretation of models

The highest values of Phoma leaf spot incidence were observed in the cooler (15–20°C) and humid periods. Other authors also observed these same conditions associated with disease progress (Pozza & Alves, 2008; Santos et al., 2014; Lorenzetti et al., 2015). Pozza and Alves (2008) reported temperatures between 16 and 20°C combined with mean daily precipitation above 4 mm to be conditions highly favorable to Phoma leaf spot. It is common to observe epidemic outbreaks of this disease in the field, at temperatures between 15 and 20°C, or even below 15°C in some regions of Brazil, according to the Phoma species that occurs in the region (Salgado & Pfenning, 2000). It is important to highlight that in our study, some incidence peaks were more associated with the occurrence of a greater number of hours between 15 and 20°C, while others were more favored by prolonged periods of leaf wetness, even with a lower occurrence of number of hours with temperatures between 15 and 20°C (data not shown).

In observations made with scanning electron microscopy, Lorenzetti et al. (2015) found that temperatures between 15 and 20°C significantly increase the germ tube length of the pathogen in a short time, thus increasing the chances of finding a wound below the cuticle that would enable infection. Additionally, low temperatures imply a reduction in the saturated vapor pressure, consequently increasing the relative humidity of the air. Therefore, even in the absence of rain, the water molecules present in the atmosphere can condense when they come into contact with the cooled leaf surface, favoring the formation of dew, which persists during the dawn and early hours of the day.

The longer duration of humidity in the microclimate favors an increase in the intensity of the disease. It is essential in the germination phase, germ tube growth, and pathogen infection. The absence of sufficient moisture in the leaf surface can affect these processes, even under ideal temperature conditions. Furthermore, the occurrence of free water on the leaf surface is responsible for lower transpiration of the crop and water potential close to zero, providing greater internal water availability for the pathogen, which may favor its colonization (Huber & Gillespie, 1992).

On the other hand, temperatures above 25°C negatively affect the fungus, in addition to reducing the duration of leaf wetness, which is essential for the Phoma leaf spot infectious process.

Lorenzetti et al. (2015) also fitted an MLR model to explain the Phoma leaf spot monocycle in coffee seedlings as a function of temperature and leaf wetness. According to the authors, temperatures between 15 and 20°C associated with periods of leaf wetness above 6 hours are favorable conditions for the disease progress, with the highest intensity observed at 48h. However, the model does not report the relationship between the number of hours in this temperature range and disease progress. Moreover, this model has no application as a warning system because it does not provide information in advance of future risks of disease progression to disease forecast.

According to Campbell and Madden (1990), warning systems should be developed using biological and environmental data to ensure reliability. Then, based on the T15-20mean, T25-30mean, RH70min, RH80max, and NDLWmax data, models 3 and 4 were able to estimate the future incidence of Phoma leaf spot in the field, two weeks in advance. This time is essential from a logistical viewpoint if fungicide spraying is necessary. The farmer in charge will need to relocate personnel and implement tasks such as spray solution preparation and application for prevention.

Validation of models

Models 3 and 4 were validated using the monthly incidence data collected in the subgrid plot. In general, the incidence values in this plot were higher than the model estimates, especially in the months of December 2013 and 2014 and in February, March, and June 2015, when the highest peaks were recorded (Figure 4). Thus, under optimal environmental conditions for disease progress, the values observed in this plot were always higher than the overall mean observed in the crop. This is because the subgrid was placed at the highest altitude location within the field, where the highest occurrence of Phoma leaf spot was observed over the years.

This information indicates the need to correct the bias of these models when implemented in software for monitoring different farms. Although they can simulate well the fluctuation of the disease over time, the mean level of disease can vary, in according to the genetics of the cultivar, altitude of the site, and crop management practices. Phoma leaf spot is a disease of economic importance in coffee crops, mainly in regions with altitude above 900 m, exactly where gourmet coffees are produced. Campbell and Madden (1990) also emphasize the importance of validating the warning system at the intended level of operation and in the region where it will be implemented to ensure it is reliable. Otherwise, the system may not meet expectations.

In fact, the results of this research are already being successfully implemented in the disease warning system, developed to predict Coffee rust and Phoma leaf spot. The forecast system is a result of the public-private partnership between the Federal University of Lavras and Cooxupé, the largest coffee cooperative in the world. Forecasts can be freely accessed for the regions where the cooperative operates, through the website: <https://sismet.cooxupe.com.br>.

Conclusion

The multiple linear regression models developed using the variables T15-20mean, T25-30mean, RH70min, RH80max, and NDLWmax, showed high precision and accuracy for modeling the Phoma leaf spot incidence in a coffee crop. Models 3 and 4 proposed in this article can may assist in decision-making regarding when to start Phoma leaf spot control with fungicides, which should occur two weeks prior to the occurrence of the disease.

Acknowledgements

We acknowledge the National Institute of Coffee Science and Technology (INCT-Café), National Council for Scientific and Technological Development (CNPq), Coordination for the Improvement of Higher Education Personnel (CAPES), and Foundation for Supporting Research of the State of Minas Gerais (FAPEMIG) for financial support. Especially the Neumann Kaffee Gruppe (NKG) farm for providing the area to perform the research study and always supporting initiatives to increase the sustainability of coffee farming.

References

- Campbell, C. L., & Madden, L. V. (1990). *Introduction to Plant Disease Epidemiology*. John Wiley & Sons.
- Companhia Nacional de Abastecimento [CONAB]. (2024). *Boletim café janeiro 2024*.
<https://www.conab.gov.br/info-agro/safras/cafe/boletim-da-safra-de-cafe>

- Huber, L., & Gillespie, T. J. (1992). Modeling leaf wetness in relation to plant disease epidemiology. *Annual Review of Phytopathology*, 30(1), 553-577. <https://doi.org/10.1146/annurev.py.30.090192.003005>
- Lorenzetti, E. R., Pozza, E. A., Souza, P. E., Santos, L. A., Alves, E., Silva, A. C., Maia, F. G. M., & Carvalho, R. C. (2015). Effect of temperature and leaf wetness on *Phoma tarda* and Phoma leaf spot in coffee seedlings. *Coffee Science*, 10(1), 1-9. <https://doi.org/10.25186/cs.v10i1.688>
- Matiello, J. B., Santinato, R., Almeida, S. R., & Garcia, A. W. R. (2015). *Cultura de Café no Brasil – Manual de Recomendações*. Fundação Procafé.
- Pozza, E. A., & Alves, M. C. (2008). Impacto potencial de mudanças climáticas sobre as doenças fúngicas do cafeeiro no Brasil. In R. Ghini, & E. Ramada (Eds.), *Mudanças climáticas: impactos sobre doenças de plantas no Brasil* (pp. 220-238). Embrapa.
- Pozza, E. A., Carvalho, V. L., & Chalfoun, S. M., (2010). Sintomas de injúrias causadas por doenças em cafeeiro. In *Semiologia do cafeeiro: sintomas de desordens nutricionais, fitossanitárias e fisiológicas*. UFLA.
- Pozza, E. A. (2021). Diagnóstico e controle de doenças. In G. R. Carvalho, A. D. Ferreira, V. T. Andrade, C. E. Botelho, & J. P. F. Carvalho (Eds.), *Cafeicultura do Cerrado*. Epamig.
- Pozza, E. A., Santos, É. R. D., Gaspar, N. A., Vilela, X. M. D. S., Alves, M. D. C., & Colares, M. R. N. (2021). Coffee rust forecast systems: Development of a warning platform in a Minas Gerais State, Brazil. *Agronomy*, 11(11), 1-22. <https://doi.org/10.3390/agronomy11112284>
- R Development Core Team. (2017). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Rossi, V., Giosuè, S., & Caffi, T. (2010). Modelling plant diseases for decision making in crop protection. In E. C. Oerke, R. Gerhards, G. Menz, & R. A. Sikora (Eds.), *Precision crop protection: The challenge and use of heterogeneity* (pp. 241-256). Springer. <https://doi.org/10.1007/978-90-481-9277-9>
- Salgado, M., & Pfenning, L. H. (2000). Identificação e caracterização morfológica de espécies de *Phoma* do cafeeiro no Brasil, In *Anais do I Simpósio de Pesquisa dos cafés do Brasil* (pp. 183-186). Embrapa.
- Santos, L. S. D., Pozza, E. A., Faria, M. A., Silva, M. L. O., Custódio, A. A. P., & Vasco, G. B. (2014). Incidência da Mancha de Phoma em cafeeiro irrigado por gotejamento, sob diferentes manejos de irrigação. *Coffee Science*, 9(1), 77-89.
- Sera, G. H., Carvalho, C. H. S., Abrahão, J. C. R., Pozza, E. A., Matiello, J. B., Almeida, S. R., Bartelega, L., & Botelho, D. M. S. (2022). Coffee leaf rust in Brazil: historical events, current situation, and control measures. *Agronomy*, 12(2), 1-19. <https://doi.org/10.3390/agronomy12020496>
- StatSoft Inc. (2004). *Statistica (data analysis software system - version 7)*.
- United States Department of Agriculture [USDA]. (2023). *Coffee: World Markets and Trade*. Foreign Agricultural Service/USDA 2023. <https://www.fas.usda.gov/data/coffee-world-markets-and-trade>