

# Forecast models of coffee leaf rust symptoms and signs based on identified microclimatic combinations in coffee-based agroforestry systems in Costa Rica

Isabelle Merle <sup>a, b, \*</sup>, Philippe Tixier <sup>c, d</sup>, Elías de Melo Virginio Filho <sup>e</sup>, Christian Cilas <sup>b, f</sup>, and Jacques Avelino <sup>a, b, e, g</sup>

<sup>a</sup> CIRAD, UPR Bioagresseurs, 30501 Turrialba, Costa Rica

<sup>b</sup> Bioagresseurs, Univ Montpellier, CIRAD, Montpellier, France

<sup>c</sup> CIRAD, UPR GECO, F-34398 Montpellier, France

<sup>d</sup> GECO, Univ Montpellier, CIRAD, Montpellier, France

<sup>e</sup> CATIE, 7170, Cartago, Turrialba, 30501, Costa Rica

<sup>f</sup> CIRAD, UPR Bioagresseurs, F-34398 Montpellier, France

<sup>g</sup> IICA AP. 55, 2200 Coronado, San José, Costa Rica

Corresponding author: E-mail: [isabelle.merle@protonmail.com](mailto:isabelle.merle@protonmail.com)

## ABSTRACT

Coffee leaf rust is a polycyclic disease that causes severe epidemics impacting yield over several years. For this reason, since the 1960s, more than 20 models have been developed to predict different indicators of the disease's development and help manage it. In existing models, standardized periods of influence of the meteorological predictors of the disease are determined *a priori*, based on strong assumptions. However, the appearance of a symptom or sign can be influenced by complex combinations of meteorological variables acting at different times and for

different durations. In our study, we monitored a total of 5400 coffee leaves during a year and a half, in different agroforestry systems, in order to detect the onset dates of the disease symptoms, such as lesion emergence, and signs, such as sporulation and infectious area increase. In these agroforestry systems, we also recorded microclimate. We statistically identified the complex combinations of microclimatic variables responsible for changes in lesion status to construct three models predicting lesion emergence probability, lesion sporulation probability and growth of its infectious area. Our method allowed the identification of different microclimatic variables that fit well with the knowledge about the coffee leaf rust biology. Minimum air temperature from 20 to 18 days before a lesion emergence explained the status change from healthy to emergence of visible lesion, possibly because the short germination phase is stimulated by low temperatures. We also found a unimodal effect of rainfall over a period of 10 days, 33 days before lesion emergence, with a maximum at 10 mm. Below this threshold, uredospore dispersal is efficient, increasing the lesion appearance probability; above this threshold, wash-off effects on uredospores probably occurs, decreasing the probability of lesion emergence. In addition, we identified microclimatic variables whose influence on coffee leaf rust had not been described before. These variables are likely to be involved in the internal development phases of the disease in the coffee leaves: (1) unimodal effects of maximum air temperature in different periods on sporulation and infectious area growth (2) positive and unimodal effects of rainfall in different periods on sporulation and (3) a negative effect of leaf thermal amplitude in different periods on lesion emergence, sporulation and infectious area growth. Although these models do not provide predictors of the level of disease attack, such as incidence, they provide valuable information for warning systems and for mechanistic model development. These models could also be used to forecast risks of infection, sporulation and infectious area growth and help optimize treatment recommendations.

Keywords: *Hemileia vastatrix*, *Coffea arabica*, weather-based statistical modeling, daily temperature, daily rainfall

## 1. Introduction

The anticipation of disease outbreaks requires knowledge of the drivers that trigger them. This knowledge can then be used to develop decision-support tools to assist farmers and thus reduce risks (Krause and Massie, 1975). Forecasting the risk of an epidemic is even more important if this epidemic leads to economic and social crises due to multiyear repercussions, as in the case of perennial crops (Cerdeira et al., 2017) and if producers are highly vulnerable. This was the case in Central America in 2012 when a severe epidemic of coffee leaf rust caused the loss of about 20% of Arabica coffee production and resulted in the declaration of a state of emergency in three countries in the region (Avelino et al., 2015). In response to this crisis, immediate measures for damage assessment, plantation rehabilitation and stakeholder training were implemented in 2013 as part of a regional project led by the Instituto Interamericano de Cooperación para la Agricultura (IICA) and Programa Cooperativo Regional para el Desarrollo Tecnológico y Modernización de la Caficultura (PROMECAFE) in collaboration with coffee-related institutes in the region. This project was financially supported by the European Union from 2017 through the Central American Program for the Integral Management of Coffee Rust (PROCAGICA), having among its objectives the development of an early warning system for the region based on surveillance and coffee leaf rust forecasting. Despite the influence of economic and social factors on plant disease outbreaks (Almeida, 2018), meteorological anomalies were considered as one of the main triggering factors of the 2012 coffee leaf rust epidemic (Avelino et al., 2015). However,

predicting disease development remains a challenge, even when focusing on meteorological variables (Cunniffe et al., 2015), and different modeling approaches can be used to develop predictive models. These approaches are often grouped into three categories: statistical models, mechanistic models and machine learning models (Siettos and Russo, 2013). The choice of an approach depends on the purpose of the model, the knowledge available on the disease and disease detection difficulties.

Statistical models depend on database analysis to describe the relationships between environmental conditions and the development of the disease characterized by symptoms, visible disease effects on the plant such as a chlorosis, and signs, physical evidence of the pathogen such as spores. These models thus have a domain of definition and are not easily generalizable. Mechanistic models have an explanatory purpose since they use equations describing the different stages of the disease's development (de Wolf and Isard, 2007). These models are therefore generalizable but require a significant amount of knowledge on the biology of the pathogen and substantial work for parameterization purposes. Finally, machine learning models make it possible to take a larger number of variables and their combinations into account for prediction but must be developed based on a learning process with a very large database. Most machine learning algorithms are considered as so-called "black box" systems, too complex to provide a biological explanation of the effects of predictors (Rudin 2019) and often limiting the user's ability to understand the results obtained (Krause and Massie, 1975).

Another important difference between the mechanistic approach and the other approaches for developing weather-based models is that statistical models and machine learning models require assumptions on the periods of action of these meteorological variables. In the field, monitoring the development of plant diseases is based on the observation of symptoms and signs and the calculation of indicators such as incidence and severity, which depend on both the development

95 of the disease and its host (Ferrandino, 2008). These variables are the result of several phases of  
96 the pathogen's development cycle. For example, for a coffee rust lesion to appear on a leaf, the  
97 fungus has to complete uredospore dispersal, deposition on the upper surface of the leaf,  
98 migration to the under surface, germination, germinative tube growth, apressorium formation to  
99 enter the leaf via a stomata and tissue colonization. Since these phases have varying durations  
100 and occur at different times, the influencing meteorological variables should affect disease at  
101 different times and during varying durations. This is even truer when considering disease  
102 descriptors such as incidence and severity that include the host plant dynamic, itself under the  
103 influence of meteorological variables.

104 In the case of coffee leaf rust, the first proposed models were multiple linear regressions  
105 between the average of minimum and maximum temperatures over incubation times (Kushalappa  
106 and Martins, 1980) or disease latency (Kushalappa and Martins, 1980; Santacreo et al., 1983;  
107 Tronconi et al., 1995). Because these experiments were based on inoculations, it might be logical  
108 not to consider variables such as rain or humidity driving uredospore dispersal, deposition and  
109 germination phases. These equations were simple but different from one site to another,  
110 indicating that they were not generalizable. These equations also considered linear effects of  
111 temperature, while controlled studies had already shown several unimodal effects of temperature  
112 on disease development (Kushalappa et al., 1983; Nutman et al., 1963). Subsequently, the models  
113 built incorporated a wide range of meteorological variables, sometimes comparing the influence  
114 of several periods (Alfonsi et al., 1974; Kushalappa, 1981; Pinto et al., 2002) but not carrying out  
115 a complete exploration of all the possible periods (Table 1). However, these studies had revealed  
116 the importance of choosing the periods of influence of meteorological variables. In addition,  
117 variables characterizing the host and quantifying the inoculum stock were considered as a way  
118 for model improvement. Further, to improve the models, incidence curves were linearized in

order to estimate the apparent infection rate of coffee leaf rust and then study the influence of weather conditions on this rate (Pedro, 1983). In those years, the most advanced model was built by Kushalappa et al. (1983, 1984). These authors developed a semi-mechanistic model in which the necessary conditions for the development of monocyclic processes were quantified and summarized in a single synthetic variable to explain the growth rate of the disease. The final modeling step was still a regression with unimodal effects of the dependent synthetic variable on the growth rate. The models developed produced recommendations regarding the appropriate dates of fungicide sprays (Kushalappa et al., 1986). The implementation of this modeling approach in Mexico led to different models from those obtained in Brazil, indicating once again their instability (Holguín, 1987).

Since 1991, most studies have focused on predicting the incidence of the disease or the infection risk rather than variables reflecting the different phases of rust development (Table 1). Since the 2000s, most models have been generated using decision trees (Avelino et al., 2006) and machine learning methods such as the support vector machine (Luaces et al., 2010), neural networks (Pinto et al., 2002), decision trees (Meira et al., 2008, 2009), fuzzy decision trees (Cintra et al., 2011) and Bayesian networks (Perez-Ariza et al., 2012). These modeling approaches have enabled inclusion of a larger number of variables (Table 1), particularly cropping practices such as fertilization and fungicide sprays (Corrales et al., 2016, 2015). However, in all cases, the periods during which the meteorological variables were considered to explain the disease were defined *a priori*, based on strong assumptions: monthly averages of meteorological variables were chosen that did not reflect the short duration of certain stages of the fungus's development. One recent study revisited mechanistic approaches, inspired by survival analysis through calculation of instantaneous risk averages as a function of meteorological variables (Bebber et al., 2016). Another recent study focusing on coffee leaf rust

prediction used the same concept of linearization of coffee leaf rust incidence used by Kushalappa in 1981 and Pedro in 1983 to study the effect of three possible periods of influence of meteorological variables (Hinnah et al., 2018) (Table 1).

We propose that improving the accuracy of prediction models does not necessarily require taking more variables into account but rather identifying the precise combinations of meteorological variables that cause onset of symptoms and signs. This approach would help us arrive at a model with a balance between the simplicity to facilitate its use and the complexity to provide sufficient precision. For coffee-based agroforestry systems, another point of improvement in prediction accuracy is to understand how shade modifies the microclimatic drivers of coffee leaf rust development, as meteorological variables are monitored by stations in full sunlight. Few of the studies cited have considered the effect of shading in their model (Avelino et al., 2006; Corrales et al., 2016, 2015; de Moraes et al., 1976). This is certainly due to the major modernization of coffee plantations after the 1970s that led to the conversion of many diversified agroforestry systems into monoculture systems, as in Brazil (Jha et al., 2014). However, agroforestry is considered as a necessary practice in the future to cope with climate change: it buffers extreme temperatures and reduces soil moisture fluctuations, for example (Lin et al., 2007). Better understanding shade effects on microclimate could be useful to select appropriate shade trees that help cope with climate change and regulate coffee leaf rust, if the critical microclimatic variables and periods of influence of the microclimate were determined.

In our study, we tried to determine, without using *a priori* assumptions, which combinations of microclimatic variables are responsible for the onset of coffee leaf rust symptoms and signs. We applied a method developed in Bugaud et al. (2015), inspired by the Window Pane approach introduced by Coakley and Line (1982), and already used to find the microclimatic periods that influence abundance of banana thrips (Carval et al., 2015), pineapple acidity (Dorey et al., 2016)

or the onset of symptoms and signs of cocoa moniliasis (Leandro-Muñoz et al., 2017). These identified microclimatic variables were used to build three models: risks of lesion occurrence, sporulation and infectious area growth. The separation into several models had the benefit of simplifying the interpretation of the microclimatic combinations identified as well as predicting specific risks that imply different recommendations in terms of chemical control.

## **2. Materials and methods**

### **2.1. Field locations**

To observe a large range of microclimatic conditions, the trial was set up in coffee plots located in experimental sites distributed in a gradient of three altitudes over a 15-month period, from May 2017 to July 2018. The first site was the long-term trial of coffee-based agroforestry systems established by the Tropical Agricultural Research and Higher Education Center (CATIE, Centro Agronómico Tropical de Investigación y Enseñanza) in Turrialba, Costa Rica (9°53'44'' latitude north, 83°40'7'' longitude west) at an altitude of 600 m.a.s.l. The second site was a coffee-based agroforestry plot established by the Costa Rican Coffee Institute (ICAFFE, Instituto del Café de Costa Rica) in Barva de Heredia, Costa Rica (10°2'9'' latitude north, 84°8'11'' longitude west) at an altitude of 1180 m.a.s.l. The third site was located in the plantation of a coffee grower living in Aserri (9°50'54'' latitude north, 84°6'0'' longitude west), at an altitude of 1500 m.a.s.l., that trial beginning in September 2017, four months later than the others.

### **2.2. Experimental design**



The trial was carried out on four plots: two in Turrialba, one in Heredia and one in Aserrí. The two selected plots in Turrialba had the same agronomic management except for the agroforestry practice, which offered different microclimatic conditions: one exposed to full sunlight; the other, in an agroforestry system based on poró (*Erythrina poeppigiana*). The plots located in Heredia and Aserrí were also agroforestry systems, based on *E. poeppigiana* in Heredia and *E. poeppigiana* combined with trees of the citrus family in Aserrí. In each plot, the coffee variety planted was susceptible to most rust races and was managed without fungicides to monitor natural behavior of coffee leaf rust. Each plot consisted of a minimum of six rows of coffee plants, with 2 m between rows and a dozen coffee plants per row, with 1-m spacing. To avoid a potential border effect, we selected the six central plants of the three central rows of each plot to monitor the coffee leaf rust, totaling 18 coffee plants per plot. In January 2018, in the plots located in Turrialba and Heredia, most of the selected plants were exhausted and deteriorated by coffee leaf rust so were replaced by new ones. Since they were selected later, in September 2017, the coffee plants of the plot in Aserrí were conserved for the 2018 monitoring. As shown in Fig. 1, the Heredia and Aserrí sites, under Pacific influence, faced a dry season from January to April 2018, while Turrialba site, under Caribbean influence, did not experience a significant dry season during the experiment. On average, higher elevation sites were those where the lowest daily minimum and maximum temperatures were measured.

### **2.3. Field monitoring of coffee leaf rust symptoms and signs and inoculum stock assessment**

Every month, three branches per coffee plant (one branch in each of three coffee plant strata) were selected and, from each branch, a node of young healthy leaves was labeled. This helped us renew the stock of healthy leaves periodically. These selected pairs of leaves were then observed

weekly over a two-month period and photographed weekly from the onset of the first visible symptom of the disease. The photographs were then analyzed by image processing with the ImageJ software (Schindelin et al., 2015) (Fig. 2). Thanks to image processing analysis, we were able to measure the lesion surfaces (symptoms and signs), from a size of 0.001 cm<sup>2</sup>. This method allowed a nondestructive monitoring of lesions over time and thus to record the history of each lesion: the week it was detected, the week the first uredospores appeared, the size of its infectious area each week. Finally, focusing on lesions rather than measuring indicators such as incidence and severity allowed us to avoid the effects of host growth dynamics on the disease. Because previous studies have shown that coffee fruit load can affect the leaf physiological resistance to the fungus (Esques and Souza 1981), we counted the number of fruiting nodes of each coffee plant studied.

Each month, to estimate the inoculum amount present in the plot, we selected 18 branches different from those used for coffee leaf rust symptoms and signs: one per selected coffee plant and six branches per foliar stratum. On these branches, every leaf with coffee leaf rust lesions showing uredospores was photographed. The rust area with uredospores was measured using the ImageJ software. The total area per branch was calculated and averaged over the 18 branches selected in each plot. This average was considered as a proxy for inoculum stock (Kushalappa, 1981; Merle et al., 2019).

#### **2.4. Microclimatic data recording**

Microclimate data were recorded using Campbell CR1000 (Campbell Scientific) data loggers placed in the center of each of the four experimental plots. Each station was equipped with a data logger connected to the following nine sensors:

- a rain gauge placed 2 m high and far from shade trees (TE525MM, accuracy 0.1 mm),
- an air temperature and relative humidity sensor positioned at 1.5 m high (HMP45C),
- four leaf wetness sensors at 1.2 m high and oriented in four different directions (Dielectric LWS),
- three T-type thermocouples (copper/constantan) giving the average temperature of three leaves each, chosen in three coffee plants and three different strata in each coffee plant (Miller, 1971).

Climatic data were recorded every five seconds by the sensors and their minimum, maximum and average values were stored every 30 minutes. Data were retrieved weekly from the data loggers using PC200W 4.5 Datalogger Support Software (Campbell Scientific).

## 2.5. Variables description

A total of 5400 leaves were monitored over all the 2-month periods, from May 2017 to July 2018. Among the total of appeared coffee leaf rust lesions, 95% ranked from the first to 25th lesion on a leaf. We therefore considered that 25 sites per leaf were available for coffee leaf rust lesion emergence. On each of the monitoring dates, each new lesion implied a change in status for a site from “healthy” to “infected.” Once the lesion appeared, this site was no longer available on the following dates because it could not become infected again. The variable presence or absence of a new coffee leaf rust lesion per foliar site (*NewLesion*) was used as a response variable to build a model forecasting the lesion emergence probability per foliar site (Fig. 2). That means that at leaf scale, the predicted number of new lesions emerged is equal to this probability multiplied by the 25 foliar sites. The date of emergence of each lesion was used to compute the lesion’s age, indicated in days but with a weekly accuracy due to the monitoring frequency

(*LesionAge*). At each monitoring date, a lesion with no uredospores had the possibility to change its status from “infected” to “infectious.” The variable presence or absence of uredospores (*Sporulation*) on the lesions was used to build a model forecasting the probability of uredospore emergence (Fig. 2). The last model aimed to predict the infectious area (*InfectiousArea*) of the lesions. In order to reduce the error about the date of site status change, we considered that this date was the middle date between the date when the change was observed and the previous monitoring date.

These three variables describing coffee leaf rust development were explained as a function of these basic daily microclimatic variables: air minimum temperature, maximum temperature and thermal amplitude (*TaMin*, *TaMax*, *TaAmp*, respectively) as well as coffee leaf minimum temperature, maximum temperature and thermal amplitude (*TcMin*, *TcMax* and *TcAmp* respectively), relative humidity amplitude (*RHAmp*), total rainfall (*Rainfall*), rainfall duration (*FreqRain*) and leaf wetness duration from 6 a.m. to 11 a.m. (*FreqLW6to11*) and from noon to 6 p.m. (*FreqLW12to18*) (Table 2). By night, the leaves were always wet in our study sites because of dew. That is why we did not consider leaf wetness duration during the night. In addition, we chose to study the morning and the afternoon separately for leaf wetness to better consider the preinfectious processes: germination usually starts in the late afternoon and continues during night, and appressorium and penetration into the leaf occur soon in the morning (Rayner, 1961). Additional variables were measured because of their known effect on coffee leaf rust development: the average inoculum stock per branch at plot level, estimated weekly by linear interpolation from monthly data (*Inoculum*), the age of the lesions (*LesionAge*), the coffee leaf stratum that marked leaves belonged to (*LeafStratum*) and the number of fruiting nodes per coffee plant (*PlantFruitLoad*) (Table 2).

## 2.6. Statistical analysis

We performed a two-step analysis. For each microclimatic variable (*MicroclimVar*), we first needed to identify which periods better explained each of the three studied dependent variables: *NewLesion*, *Sporulation* and *InfectiousArea*. A period is defined by a duration (D) and a time before status change (T) (from a healthy foliar site to the emergence of a lesion, from a latent lesion to an infectious lesion, from the past infectious area to the current infectious area). The method used described by Bugaud et al. in 2015, proceeds as follows: (1) we calculated, over all the possible periods (T ranges from 1 to n and, for each time T, D ranges from 1 to T), the explanatory microclimatic variables  $MicroclimVar_{TxDy}$ , defined as the daily averaged microclimatic variable (Table 2) over a y-days period starting x days before the symptom or sign onset (2) each of these explanatory variables was included in a generalized linear model (GLM) to explain *NewLesion*, *Sporulation* and *InfectiousArea* with different distributions depending on the dependent variable (see Eqs. 1, 2 and 3); for rainfall and temperatures, which generally have a nonlinear effect, we tested their unimodal form; (3) explanation level provided by the variables was then assessed by the difference between the Akaike information criterion (AIC) of the model with the microclimatic variable and the AIC of a specific reference model without this variable (see following equations); AIC is an indicator of “relative goodness of fit of a statistical model for a given data set” (Akaike, 1974); (4) the variable whose model maximized the difference, i.e. the most explanatory variable, was retained in the second step. Considering the second selection criterion of the Window Pane approach mentioned by Pietravalle et al. (2003), consecutive periods, i.e. same duration but a starting day differing by one day or same starting day but a duration differing by one day should result in similar AIC due to autocorrelations between these

periods. The last selection criterion (Pietravalle et al., 2003) was the similarity of the selected periods between related microclimatic variables such as leaf and air temperatures.

$$NewLesion \sim PastInoculum + MicroclimVar_{Tx Dy} \quad (binomial GLM) \quad (1)$$

$$reference model: NewLesion \sim PastInoculum \quad (binomial GLM)$$

$$Sporulation \sim LesionAge + MicroclimVar_{Tx Dy} \quad (binomial GLM) \quad (2)$$

$$reference model: Sporulation \sim LesionAge \quad (binomial GLM)$$

$$InfectiousArea \sim InfectiousArea_{day-7} + MicroclimVar_{Tx Dy} \quad (gaussian GLM) \quad (3)$$

$$reference model: InfectiousArea \sim InfectiousArea_{day-7} \quad (gaussian GLM)$$

We included a variant in this general method by incorporating in the models some variables that necessarily explain *NewLesion*, *Sporulation* or *InfectiousArea*. Hence, in the models for explaining *NewLesion*, we incorporated a variable called *PastInoculum* (see Eq. 1). This variable represents the averaged inoculum stock available in a specific period before the observation of new coffee leaf rust lesions. This inclusion seemed appropriate as no infection can occur without inoculum. We did not choose an arbitrary period. On the contrary, we used the same method previously described to identify the best past inoculum period to explain *NewLesion* (see Eq. 4). In the models used, we considered the coffee leaf stratum because incidence and severity, and hence quantity of uredospores, have been reported to be higher in the low coffee tree strata (Avelino et al. 1991; Villegas-García and Baeza-Aragón, 1990).

$$NewLesion \sim LeafStratum + Inoculum_{period} \quad (binomial GLM) \quad (4)$$

$$reference model: NewLesion \sim LeafStratum \quad (binomial GLM)$$

Similarly, for the sporulation model, the age of the lesion (*LesionAge*) entered in the model since an older lesion has a higher probability of sporulating than a younger one (see Eq. 2). In the last model, for *InfectiousArea* modeling, we included the past infectious area of the lesion (one week before): *InfectiousArea*<sub>day-7</sub> (see Eq. 3). This variable is the lesion infectious area (*InfectiousArea*) measured at the previous monitoring date, so it is a reference for infectious area growth.

The second step consisted of building complete generalized linear models (GLMs) for *NewLesion*, *Sporulation* and *InfectiousArea* by including the best microclimatic variables identified in Step One. Only uncorrelated predictors were included in the models. For that purpose, we analyzed the correlation between variables and discarded those with  $R^2 > 0.49$  (Dormann et al., 2012). We also used biological criteria to finalize certain choices for several variables. The variable *PlantFruitLoad* was included in all of the models for its possible effects on coffee leaf rust development (López-Bravo et al., 2012).

Prior to proceeding with the complete model step, the distributions of the selected variables were studied to focus on domains of definition with a sufficient number of observations. When there were interactions between variables, a division into submodels was performed using the "party" package (Hothorn et al., 2006), which builds a tree-based regression by recursive binary partitioning. The binomial models were then evaluated using the package pROC (Robin et al., 2011) that transforms the response variable into a binary response, finding the threshold that maximizes the area under the receiver operating characteristic (ROC) curve. The classification error is then determined by the confusion matrix. We evaluated the Gaussian models by building the regression between observed and predicted data. All statistical analyses were performed with R 3.5.1 (R Development Core Team 2018) and with an alpha level of 0.05. Eqs. 1, 2, 3 and 4 and their reference models (Step One) and complete models (Step Two) were fitted with the GLM

function in the lme4 package (Bates et al., 2015). The maximum likelihood of parameters was approximated by the Laplace method (Bolker et al., 2009).

### 3. Results

In the graphs generated in Step One of the analysis, the largest differences in AIC indicate a strong influence of the microclimatic variable averaged on the associated period (Figs. 3, 4, 5, 6). Due to the second selection criteria (Pietravalle et al. 2003) and our approximation of symptom or sign onset by the midpoint of the two monitoring dates, the periods of influence were visible on the graphs by areas of higher AICs rather than punctual dots. Table 3 indicates the selected periods for each microclimatic variable: a starting date for the consideration of the variable (T), in days before foliar site status change, and duration in days (D). Only uncorrelated and significant variables were kept in the models. In the final models equations, we indicated the parameter estimates with Greek letters described in Table 4.

#### 3.1. Predictive model of the probability of lesion emergence per foliar site

By exploring the influence periods of the inoculum stock on *NewLesion*, we were able to identify the best period of influence of *PastInoculum* (Eq. 1): between five and three weeks before the lesion appeared (Fig. 3). From the 26 microclimatic variables resulting from this first selection (Fig. 4), only three uncorrelated variables were conserved and had a significant effect: the average daily rainfall recorded between 33 and 24 days before lesion emergence ( $Rain_{T33D10}$ ), the average daily minimum leaf temperature recorded between 20 and 18 days before lesion emergence ( $TcMin_{T20D3}$ ) and the average daily thermal amplitude of the leaves recorded between



20 and 10 days before lesion emergence ( $TcAmp_{T20D11}$ ). Prior to design of the complete model, we removed the microclimatic data combinations not well-represented in the distributions of these variables (Fig. 7). The final data sets of models A1 and A2 contained 223 945 and 347 491 observations respectively. We found interactions including the average rainfall recorded between 33 and 24 days before lesion emergence ( $Rain_{T33D10}$ ). Because this variable had a unimodal effect, we created two models (A1 and A2) based on two rainfall classes determined by the tree-based regression built by recursive binary partitioning:  $\leq 4$  mm per day and  $>4$  mm per day.

Model A1 ( $n = 223945, Rain_{T33D10} \leq 4$  mm): with  $x = \alpha_{1.1}$

$$NewLesion = \frac{\exp(x)}{1 + \exp(x)} + \alpha_{1.2} \times PastInoculum + \alpha_{1.3} \times Rain_{T33D10} + \alpha_{1.4} \times TcMin_{T20D3} + \alpha_{1.5} \times TcAmp_{T20D11}$$

Model A2 ( $n = 347491, Rain_{T33D10} > 4$  mm): with  $x = \alpha_{2.1}$

$$NewLesion = \frac{\exp(x)}{1 + \exp(x)} + \alpha_{2.2} \times PastInoculum + \alpha_{2.3} \times Rain_{T33D10} + \alpha_{2.4} \times (Rain_{T33D10})^2 + \alpha_{2.5} \times TcMin_{T20D3} + \alpha_{2.6} \times (TcMin_{T20D3})^2 + \alpha_{2.7} \times TcAmp_{T20D11}$$

In both models, *PastInoculum* showed a positive effect on *NewLesion* ( $p < 0.0001$ ). In model A1, the effect of the  $Rain_{T33D10}$  variable was positive ( $p < 0.0001$ ) (Fig. 8A), while its effect was unimodal with a maximum around 10 mm in model A2 ( $p < 0.0001$ ) (Fig. 8B).  $TcMin_{T20D3}$  had a positive effect on *NewLesion* in model A1 ( $p < 0.0001$ ) and a unimodal effect with a maximum of

about 19.5°C in model A2 ( $p < 0.0001$ ).  $TcAmp_{T20D11}$  negatively affected *NewLesion* in both models ( $p < 0.0486$  in model A1 and  $p < 0.0001$  in model A2). Predicted values of *NewLesion* as a function of  $Rain_{T33D10}$ ,  $TcMin_{T20D3}$ ,  $TcAmp_{T20D11}$  and *PastInoculum* within the range of observed values are presented in Fig. 8 (A1 and A2 graphs). After transformation into binary responses using classification thresholds of 0.0141 for the A1 model and 0.0590 for A2 model, the confusion matrix of models A1 and A2 indicated a classification error of 0.32 and 0.24, respectively. The *PlantFruitLoad* variable was not kept in the complete model because it did not improve the classification error of each of these models. Models A1 and A2 together expressed a root mean square error of 0.032 in predicting the lesion emergence probability, by plot and by date (Fig. 9A).

### 3.2. Predictive model of sporulation probability of a coffee leaf rust lesion

We selected 23 microclimatic variables associated with a period (Fig. 5), resulting from Step One of our analysis method, as the best predictors of *Sporulation*. Among these variables, only seven were conserved after removing the highly correlated ones (Table 3). The complete model included four main microclimatic variables: the average daily maximum air temperature recorded between 15 and 12 days before lesion sporulation ( $TaMax_{T15D4}$ ), the average daily rainfall recorded between 12 and 11 days before lesion sporulation ( $Rain_{T12D2}$ ), the average daily rainfall recorded between five and three days before lesion sporulation ( $Rain_{T5D3}$ ) and the average daily thermal amplitude of the leaves recorded four days before lesion sporulation ( $TcAmp_{T4D1}$ ). Among the seven variables that influenced *Sporulation*, two variables were related to the pathogen and the host plant: *LesionAge* and *PlantFruitLoad*.

Model B ( $n = 51539$ ):

$$\begin{aligned} \text{Sporulation} = \frac{\exp(x)}{1 + \exp(x)} \quad \text{with } x = & \beta_1 \\ & + \beta_2 \times \text{LesionAge} + \beta_3 \times (\text{LesionAge})^2 \\ & + \beta_4 \times \text{PlantFruitLoad} \\ & + \beta_5 \times \text{TaMax}_{T15D4} + \beta_6 \times (\text{TaMax}_{T15D4})^2 \\ & + \beta_7 \times \text{Rain}_{T12D2} \\ & + \beta_8 \times \text{Rain}_{T5D3} + \beta_9 \times (\text{Rain}_{T5D3})^2 \\ & + \beta_{10} \times \text{TcAmp}_{T4D1} \end{aligned}$$

In order to focus on a domain of definition with a sufficient number of observations, we first removed the microclimatic data combinations not well represented in the distributions, as described in Fig. 7 for models A1 and A2. The final data set of model B contained 51 539 observations. *LesionAge* was the variable that most influenced *Sporulation* (Model B), with a maximum effect about 15 days before the date of foliar site status change ( $p < 0.0001$ ). *PlantFruitLoad*, which characterizes host phenology, positively influenced *Sporulation* ( $p = 0.0005$ ). The microclimatic variables *TaMax*<sub>T15D4</sub> and *Rain*<sub>T5D3</sub> showed a significant unimodal effect ( $p < 0.0001$ ) on *Sporulation*, with respective maxima about 28°C and 10 mm (Fig. 8C). In addition, *Rain*<sub>T12D2</sub> had a positive effect on *Sporulation* ( $p < 0.0001$ ), while *TcAmp*<sub>T4D1</sub> affected it negatively ( $p < 0.0001$ ). Predicted values of *Sporulation* as a function of *TaMax*<sub>T15D4</sub>, *Rain*<sub>T5D3</sub>, *Rain*<sub>T12D2</sub>, *TcAmp*<sub>T4D1</sub>, *LesionAge* and *PlantFruitLoad*, within the range of observed values, are presented in Fig. 8B. After transformation into binary responses using a classification threshold of 0.140, the classification error computed by the confusion matrix was 0.39 for this model. However, Model B expressed a root mean square error of 0.070 in predicting the sporulation probability, by plot and by date (Fig. 9B).

### 3.3. Model forecasting coffee leaf rust lesion infectious area

For this last model, eight microclimatic variables were selected at the end of Step One (Fig. 6). Of these variables, only three were kept after removing the highly correlated variables and only two had a significant effect in the complete model (Table 3).

$$\begin{aligned} \text{Model C } (n = 16363): \quad \text{InfectiousArea} = & + \gamma_1 \\ & + \gamma_2 \times \text{InfectiousArea}_{\text{day}-7} \\ & + \gamma_3 \times \text{TaMax}_{T5D1} + \gamma_4 \times (\text{TaMax}_{T5D1})^2 \\ & + \gamma_5 \times \text{TcAmp}_{T1D1} \end{aligned}$$

*InfectiousArea* was mainly explained by *InfectiousArea*<sub>day-7</sub>. *TaMax*<sub>T5D1</sub> had a unimodal effect, with a maximum of about 28°C (p < 0.0001) (Fig. 8D). *InfectiousArea* was also affected negatively by *TcAmp*<sub>T1D1</sub> (p < 0.0001). *PlantFruitLoad* had no significant effect. As we did for models A1 and A2, as shown in Fig. 7, we selected domains of definition of *TaMax*<sub>T5D1</sub> and *TcAmp*<sub>T1D1</sub> in order to include combinations that represented the data well in model C. These definition domains are shown in Fig. 8C. The final data set of model C contained 16 363 observations. Since this model had a Gaussian response, we did not use the confusion matrix. Model C expressed a root mean square error of 0.012 in predicting the lesion infectious area, by plot and by date (Fig. 9C).

## 4. Discussion

By modeling foliar site status changes, i.e. lesion level changes, we were able to determine, with no assumptions, complex combinations of microclimatic variables that influence several disease processes. In addition, this approach helped us isolate disease growth from the host

dynamics, which is not possible when studying synthetic disease descriptors as incidence or severity (Ferrandino, 2008). As already demonstrated by other studies (Bugaud et al., 2015; Carval et al., 2015; Dorey et al., 2016; Leandro-Muñoz et al., 2017), our results show the importance of exploring different periods of influence of microclimatic variables, as these can vary according to the considered variable and their specific effect. Short duration periods, some only days long, have been identified for several microclimatic variables. These would have been masked if we had considered long periods determined in advance. We showed that the periods of influence of microclimatic variables follow on logically according to their specific effect on disease development instead of overlapping (Fig. 10). Not all of the favourable conditions for disease growth need to be met at the same time and during the same period.

Our two models on lesion emergence probability incorporated more adjusted microclimatic variables than the empirical models obtained so far. These variables are consistent with the knowledge of the factors affecting the biology of the fungus. Based on the best estimated inoculum predictor for lesion emergence probability — the period from five to three weeks before the lesion emergence — the incubation period would be at least 21 days. In addition, the period from 20 to 18 days before a lesion emergence in which an influence of the minimum leaf temperature was observed could be related to the uredospore germination and penetration. Indeed, germination is stimulated by the previous exposition to low temperatures recorded at night, and the duration of three days could evoke a survival time of the uredospores (Nutman et al., 1963). This hypothesis would imply that the incubation period would be around three weeks, a result in agreement with the findings of Leguizamón-Caycedo et al. (1998), who measured incubation periods ranging from 18 to 24 days depending on the level of shade in their studies on the Caturra coffee variety in Colombia.

In our study, the unimodal effect of the average rainfall between 33 to 24 days before lesion emergence could therefore correspond to the dissemination phase of the uredospores. We hypothesize that uredospore release and deposition occur with rainfall below 10 mm, in accordance with previous works (Bock, 1962; Nutman et al., 1960). However, above 10 mm, uredospore wash-off by rain probably occurs, decreasing the probability of lesion emergence. Uredospore wash-off by rain has already been recognized as a strong limitation for *H. vastatrix* propagation (Avelino et al., 2019; Savary et al., 2004) and in other pathosystems (Sache, 2000). Under the conditions of our trial, in agroforestry systems, dispersal is mainly carried out by rain through raindrop impact on leaves, rather than by wind (Boudrot et al., 2016). The presence of trees in agroforestry systems and in the landscape helps intercept wind, limiting its effect (Avelino et al., 2012; Boudrot et al., 2016; Pezzopane et al., 2011). In other conditions, however, wind has been considered as the main factor affecting uredospore dispersal (Becker et al., 1975).

Once the uredospore is located on the underside surface of the leaf, where the stomata that serve as entry doors to the fungus are located, uredospore germination, germ tube growth and appressorium formation phases require the presence of water and low light intensity (Rayner 1961). In our study, the water required for these phases was not a limiting factor since the leaves were always wet at night because of dew. The last microclimatic variable included in our two models on lesion emergence probability was the average thermal amplitude of the leaves between 20 and 10 days before symptom onset. This finding possibly illustrates the influence of leaf temperature on the colonization phase of the fungus (Kushalappa and Eskes, 1989; Ribeiro, 1978). The values of the lesion emergence probability are in general low: 75% of the values below a probability of 0.04 (Fig. 9A). However, as we subdivided the coffee leaf into 25 foliar sites for potential lesion emergence, this model calculates a probability of occurrence of a lesion per site and not per leaf. A probability of 0.04 at the foliar site scale represents a probability 25

times higher at the leaf scale, therefore a probability equal to one for a leaf to become diseased and express the symptoms.

Given the complexity of investigating internal leaf-development phases (McCain and Hennen, 1984), most studies have focused on studying the latency period. In the case of coffee leaf rust, this period includes the incubation period and the time required by the pathogen to colonize the leaf and produce its first infectious entities via stomata. In our predictive model on sporulation, the most important variable was the age of the lesion, with an optimal age of 15 days to sporulate (Fig. 8C). Microclimatic variables were secondary. This result indicates that once into the leaf, the pathogen was less dependent on external conditions for its growth and sporulation. The optimal age of a lesion for sporulation is the time that the pathogen requires to colonize the mesophyll sufficiently and to accumulate in the substomatal spaces from which the uredospore precursor cells will be differentiated and released (McCain and Hennen, 1984). This optimal lesion age can also be the time needed to accumulate the necessary nutrients for the development of these cells.

Few studies have investigated the microclimatic factors that influence the processes of sporulation and growth of the lesion infectious area independently from the infection phase. We hypothesized that the microclimatic periods that we identified were related with stomata functioning that, in turn, affected the occurrence of the subprocesses of coffee leaf rust sporulation (Guggenheim and Harr, 1978). The positive rainfall influence between 12 and 11 days and between five and three days before lesion sporulation can be explained by the opening of the stomata under high humidity conditions (Butler, 1977). Stomata opening facilitates the release of sporogenous cells (Guggenheim and Harr, 1978), with a number of emerging sporogenous cells increased with a larger ostiole opening (McCain and Hennen, 1984). In addition, daily leaf temperature amplitude four days before lesion sporulation negatively affected

sporulation, possibly because high temperatures, increasing thermal amplitude, promote stomata closure, due to CO<sub>2</sub> accumulation in the substomatal areas (Heath and Orchard, 1957). Another possible hypothesis is the lethal effect of high temperatures on fungus internal development (Ribeiro, 1978). Finally, the unimodal effect of the daily average of the maximum air temperatures from 15 to 12 days before lesion sporulation, with an optimal temperature around 28°C, can be explained by negative effects of low temperatures (Toniutti et al., 2017) and high temperatures on sporulation. This effect could also be related to the colonization phase, as 28°C is the maximum temperature limit for the latent period (Waller, 1982). Finally, the negative part of the unimodal effect of rainfall, from five to three days before lesion sporulation, is not easy to interpret. It can be assumed that days with heavy rainfall will generally be related to lower maximum temperatures, less favorable to sporulation.

The last variable included in the sporulation model was the plant fruit load, estimated by the number of fruiting nodes per plant, which was found to have a positive effect on the sporulation probability of the lesions. This result is compatible with the hypothesis of a migration, from leaves to fruits, of phenolic compounds involved in the plant's defense mechanisms (Chalfoun and Carvalho, 1987).

Despite these interesting results, our model appeared rather inaccurate since the model explained only 26% of the sporulation probability. This low precision is possibly due to a high variability in leaf temperature within the same coffee layer (Miller, 1971) that our measurements did not capture. Indeed, the radiation received by the leaves can vary greatly, especially in the lower strata, which benefit from the partial and irregular shade provided by the upper stratum of the coffee plants and that of neighboring coffee plants (Butler, 1977). Another hypothesis to explain the inaccuracy could be our consideration of plant fruit load in the model without taking into account the development stage of fruits. Indeed, within fruit development, there is a



progressive migration from leaves to fruits of phenolic compounds known for their role in plant defense mechanisms (de Carvalho et al., 2001; Kushalappa and Eskes, 1989).

With regard to the model predicting lesion infectious area growth, the variable that logically emerged as the main factor was the past lesion infectious area measured at the last monitoring date. However, this growth was dependent on microclimate variations. Infectious area was larger when maximum air temperature was about 28°C five days before area measurement and leaf temperature amplitude the day before was lower. These short duration effects are probably related to brief phenomena such as stomata opening, releasing more sori. These effects could be similar to those of maximum air temperature and leaf temperature amplitude on sporulation onset.

It is important to focus on the predictive potential of the models in terms of time of anticipation. In all of the models we developed, some weather variables had an effect close to the onset of symptoms and signs. In the case of the models predicting the probability of lesion emergence, the earliest influencing variables, the inoculum amount between five and three weeks before lesion emergence and the average rainfall intensity between 30 and 24 days before, can be used to calculate intervals of lesion emergence probability or provide possible scenarios. In contrast, weather does not have an early influence on sporulation and infectious area growth. In these cases, the use of meteorological forecasts with a confidence interval, instead of meteorological measurements, could be used in our models to compute response intervals and suggest possible future scenarios.

Our models require monitoring of variables such as the inoculum stock, lesion age and past lesion infectious area. These variables are good predictors, respectively, of infection levels (Kushalappa, 1981), sporulation probability and lesion infectious area and significantly

contribute to the accuracy of plant disease prediction models (Krause and Massie, 1975), therefore deserving to be assessed in coffee leaf rust surveillance activities.

## 5. Conclusion

Our statistical analysis permitted determination, without *a priori* assumptions, of periods of influence per microclimatic variable that vary according to their specific effect on coffee leaf rust development stages instead of hypothetical periods of influence identical for all meteorological variables. This process-based approach, such as mechanistic models, improved our understanding of coffee leaf rust development and enabled building models forecasting the onset of coffee leaf rust symptoms and signs. It is likely that these models can be used in Central American coffee areas with meteorological variables that fit within our domains of definition to predict different risks in terms of coffee leaf rust development stages. These different development stages imply different fungicide type recommendations: protectant fungicides to prevent infection; curative fungicides to suppress colonization and sporulation. However, the accuracy of the sporulation probability model still needs to be improved to consider its application, by incorporating missing variables, as possibly fruit phenology. Due to their simplicity, these models have the advantage that they can be easily evaluated with other datasets and improved. Their use, combined with a crop model in a simulator, could even make it possible to compare the simulated incidences with incidences measured by the monitoring programs of the different Central American countries. The combinations of microclimatic variables that we determined as influencing coffee leaf rust growth also represent valuable information for the development of a mechanistic model. Finally, it would be important to study how agroforestry systems and their management can help regulate

coffee leaf rust by modifying potential favourable microclimate conditions to this disease in the understory.

## **Acknowledgements**

For their valuable technical work, we thank Alejandra Barquero, Hugo Mendez and Steven Cerdas. This work was developed as part of the “Programa Centroamericano de Gestión Integral de la Roya del Café” (PROCAGICA) funded by the EU (DCI-ALA/2015/365-17). We thank the Ernesto Illy Foundation and CIRAD (Centre de Coopération Internationale en Recherche Agronomique pour le Développement) for their financial support, CATIE, ICAFE and the coffee grower Enrique Montenegro for allowing access to their coffee-based agroforestry systems, and R.Scheck for reviewing the English.

## References

- Alfonsi, R.R., Ortolani, A.A., Pinto, H.S., Pedro Junior, M.J., Brunini, O., 1974. Associação entre nível de infecção da ferrugem do cafeeiro, variáveis climáticas e área foliar, observadas em *Coffea arabica*, in: Congresso Brasileiro Sobre Pesquisas Cafeeiras. pp. 80–83.
- Almeida, R.P.P., 2018. Emerging plant disease epidemics: Biological research is key but not enough. PLOS Biol 16, e2007020. <https://doi.org/10.1371/journal.pbio.2007020>
- Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle, in: Petrov, B.N., Csáki, F. (Eds.), 2nd International Symposium on Information Theory. Tsahkadsor, Armenia, USSR, Budapest: Akadémiai Kiadó, 267–281.
- Avelino, J., Badaroux, J., Boudrot, A., Brenes Loaiza, M.A., Granados, E., Henrion, M., Lopez, D., Merle, I., Pico Rosado, J.T., Segura, B., Vilchez Mendoza, S.J., Smith, M., De Melo, E., 2019. Shade effects on coffee rust (*Hemileia vastatrix*). In: Dupraz Christian, G.M., Lawson Gerry (Ed.), Proceedings of the 4th World Congress on Agroforestry. Book of abstracts, 05-20/05-22 2019, Montpellier, France.
- Avelino, J., Cristancho, M., Georgiou, S., Imbach, P., Aguilar, L., Bornemann, G., Läderach, P., Anzueto, F., Hruska, A.J., Morales, C., 2015. The coffee rust crises in Colombia and Central America (2008–2013): impacts, plausible causes and proposed solutions. Food Secur. 7, 303–321. <https://doi.org/10.1007/s12571-015-0446-9>
- Avelino, J., Romero-Gurdián, A., Cruz-Cuellar, H.F., Declerck, F.A., 2012. Landscape context and scale differentially impact coffee leaf rust, coffee berry borer, and coffee root-knot nematodes. Ecol. Appl. 22, 584–596.
- Avelino, J., Zelaya, H., Merlo, A., Pineda, A., Ordoñez, M., Savary, S., 2006. The intensity of a coffee rust epidemic is dependent on production situations. Ecol. Model. 197, 431–447. <https://doi.org/10.1016/j.ecolmodel.2006.03.013>
- Avelino, J., Muller, R.A., Cilas, C., Velasco Pascual, H., 1991. Development and behavior of coffee orange rust (*Hemileia vastatrix* Berk. and Br.) in plantations undergoing modernization, planted with dwarf varieties in South-East Mexico. Café Cacao Thé. 35(1), 21–37.
- Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting Linear Mixed-Effects Models Using lme4. J. Stat. Soft. 67. <https://doi.org/10.18637/jss.v067.i01>
- Bebber, D.P., Castillo, Á.D., Gurr, S.J., 2016. Modelling coffee leaf rust risk in Colombia with climate reanalysis data. Philos. Tans. R. Soc. B 371, 20150458. <https://doi.org/10.1098/rstb.2015.0458>
- Becker, S., Mulinge, S.K., Kranz, J., 1975. Evidence that uredospores of *Hemileia vastatrix* Berk. and Br. are wind-borne. Phytopathol. Z. 82, 359–360.
- Bock, K.R., 1962. Dispersal of uredospores of *Hemileia vastatrix* under field conditions. Trans. Brit. Mycol. Soc. 45, 63–74. [https://doi.org/10.1016/S0007-1536\(62\)80035-7](https://doi.org/10.1016/S0007-1536(62)80035-7)
- Bolker, B.M., Brooks, M.E., Clark, C.J., Geange, S.W., Poulsen, J.R., Stevens, M.H.H., White, J.-S.S., 2009. Generalized linear mixed models: a practical guide for ecology and evolution. Trends Ecol. Evol. 24, 127–135. <https://doi.org/10.1016/j.tree.2008.10.008>
- Boudrot, A., Pico, J., Merle, I., Granados, E., Vilchez, S., Tixier, P., Filho, E. de M.V., Casanoves, F., Tapia, A., Allinne, C., Rice, R.A., Avelino, J., 2016. Shade Effects on the Dispersal of Airborne *Hemileia vastatrix* Uredospores. Phytopathology 106, 572–580. <https://doi.org/10.1094/PHYTO-02-15-0058-R>

655 Bugaud, C., Joannès-Dumec, C., Louisor, J., Tixier, P., Salmon, F., 2015. Preharvest temperature affects  
656 chilling injury in dessert bananas during storage: Preharvest temperature affects chilling injury in  
657 dessert bananas during storage. *J. Sci. Food Agric.* 96, 2384–2390.  
658 <https://doi.org/10.1002/jsfa.7354>

659 Butler, D.R., 1977. Coffee Leaf Temperatures in a Tropical Environment. *Acta Bot. Neerl.* 26, 129–140.  
660 <https://doi.org/10.1111/j.1438-8677.1977.tb01106.x>

661 Carval, D., Cotté, V., Notaro, M., Ryckewaert, P., Tixier, P., 2015. Spatiotemporal population dynamics  
662 of the banana rind thrips, *Elixothrips brevisetis* (Bagnall) (Thysanoptera: Thripidae). *J. Appl.*  
663 *Entomol.* 139, 510–518. <https://doi.org/10.1111/jen.12190>

664 Cerda, R., Avelino, J., Gary, C., Tixier, P., Lechevallier, E., Allinne, C., 2017. Primary and secondary  
665 yield losses caused by pests and diseases: Assessment and modeling in coffee. *PLOS ONE* 12,  
666 e0169133.

667 Chalfoun, S.M., de Carvalho, V.D., 1987. Efeito da produção e da composição química de folhas de  
668 cafeeiros sobre a intensidade de ataque de ferrugem (*Hemileia vastatrix* Berk & Br.), in: 14<sup>o</sup>  
669 Congresso Brasileiro de Pesquisas Cafeeiras, 1<sup>o</sup> Congresso Latinoamericano de Tecnologia  
670 Cafeeira, Ministerio da Indústria e do Comercio, Instituto Brasileiro do Café: Campinas 121-122.

671 Cintra, M.E., Meira, C.A.A., Monard, M.C., Camargo, H.A., Rodrigues, L.H.A., 2011. The use of fuzzy  
672 decision trees for coffee rust warning in Brazilian crops, in: 2011 11th International Conference  
673 on Intelligent Systems Design and Applications. IEEE, Cordoba, Spain, pp. 1347–1352.  
674 <https://doi.org/10.1109/ISDA.2011.6121847>

675 Coakley, S.M., Line, R.F., 1982. Prediction of stripe rust epidemics on winter wheat using statistical  
676 models. *Phytopathology* 72, 1006.

677 Corrales, D.C., Casas, A.F., Ledezma, A., Corrales, J.C., 2016. Two-Level Classifier Ensembles for  
678 Coffee Rust Estimation in Colombian Crops: *International Journal of Agricultural and*  
679 *Environmental Information Systems* 7, 41–59. <https://doi.org/10.4018/IJAEIS.2016070103>

680 Corrales, D.C., Figueroa, A., Ledezma, A., Corrales, J.C., 2015. An Empirical Multi-classifier for Coffee  
681 Rust Detection in Colombian Crops, in: Gervasi, O., Murgante, B., Misra, S., Gavrilova, M.L.,  
682 Rocha, A.M.A.C., Torre, C., Taniar, D., Apduhan, B.O. (Eds.), *Computational Science and Its*  
683 *Applications -- ICCSA 2015*. Springer International Publishing, Cham, pp. 60–74.  
684 [https://doi.org/10.1007/978-3-319-21404-7\\_5](https://doi.org/10.1007/978-3-319-21404-7_5)

685 Cuniffe, N.J., Koskella, B., E. Metcalf, C.J., Parnell, S., Gottwald, T.R., Gilligan, C.A., 2015. Thirteen  
686 challenges in modelling plant diseases. *Epidemics* 10, 6–10.  
687 <https://doi.org/10.1016/j.epidem.2014.06.002>

688 de Carvalho, V.L., Chalfoun, S.M., Castro, H.A., de Carvalho, V.D., 2001. Influência de diferentes níveis  
689 de produção sobre a evolução da ferrugem do cafeeiro e sobre teores foliares de compostos  
690 fenólicos. *Ciênc. agrotec.* 25, 49–54.

691 de Moraes, S.A., Sugimori, M.H., Ribeiro, I.J.A., Ortolani, A.A., Pedro Jr., M.J., 1976. Incubation period  
692 of *Hemileia vastatrix* B. et Br. in three regions of Sao Paulo State. *Summa Phytopathol.* 2, 32–38.

693 de Wolf, E.D., Isard, S.A., 2007. Disease Cycle Approach to Plant Disease Prediction. *Annu. Rev.*  
694 *Phytopathol.* 45, 203–220. <https://doi.org/10.1146/annurev.phyto.44.070505.143329>

695 Dorey, E., Fournier, P., Léchaudel, M., Tixier, P., 2016. A statistical model to predict titratable acidity of  
696 pineapple during fruit developing period responding to climatic variables. *Scientia Horticulturae*  
697 210, 19–24. <https://doi.org/10.1016/j.scienta.2016.07.014>

698 Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Garcia Marquéz, J.R., Gruber, B.,  
699 Lafourcade, B., Leitão, P.J., Münkemüller, T., McClean, C., Osborne, P.E., Reineking, B.,  
700 Schröder, B., Skidmore, A.K., Zurell, D., Lautenbach, S., 2012. Collinearity: a review of methods  
701 to deal with it and a simulation study evaluating their performance. *Ecography* 35, 1–20.

702 Eskes, A.B., De Souza, E.Z., 1981. Ataque da ferrugem em ramos com e sem produção, de plantas do  
703 cultivar catuaí. In: Congresso Brasileiro de Pesquisas Cafeeiras, Sao Lourenço, Minas Gerais,  
704 (Brasil) : IBC, 186-8.

705 Ferrandino, F.J., 2008. Effect of Crop Growth and Canopy Filtration on the Dynamics of Plant Disease  
706 Epidemics Spread by Aerially Dispersed Spores. *Phytopathology* 98, 492–503.  
707 <https://doi.org/10.1094/PHYTO-98-5-0492>

708 Guggenheim, R., Harr, J., 1978. Contributions to the Biology of *Hemileia vastatrix* II. SEM-Investigations  
709 on Sporulation of *Hemileia vastatrix* on Leaf Surfaces of *Coffea arabica*. *J. Phytopathol.* 92, 97–  
710 101. <https://doi.org/10.1111/j.1439-0434.1978.tb03590.x>

711 Jha, S., Bacon, C.M., Philpott, S.M., Ernesto Méndez, V., Läderach, P., Rice, R.A. 2014. Shade Coffee:  
712 Update on a Disappearing Refuge for Biodiversity. *BioScience*. 64, 416–428.

713 Heath, O.V.S., Orchard, B., 1957. Midday closure of stomata. Temperature effects on the minimum  
714 intercellular space carbon dioxide concentration “T”. *Nature* 180, 180-181.

715 Hinnah, F.D., Sentelhas, P.C., Meira, C.A.A., Paiva, R.N., 2018. Weather-based coffee leaf rust apparent  
716 infection rate modeling. *Int. J. Biometeorol.* 62, 1847–1860. [https://doi.org/10.1007/s00484-018-](https://doi.org/10.1007/s00484-018-1587-2)  
717 1587-2

718 Holguín, F., 1987. Estudios epidemiológicos de la roya del cafeto en México, in: X Simposio  
719 Latinoamericano sobre Caficultura, IICA: Tapachula, Chiapas, (México). 32-39.

720 Hothorn, T., Hornik, K., Zeileis, A., 2006. Unbiased Recursive Partitioning: A Conditional Inference  
721 Framework. *J. Comput. Graph. Stat.* 15, 651–674. <https://doi.org/10.1198/106186006X133933>

722 Krause, R.A., Massie, L.B., 1975. Predictive Systems: Modern Approaches to Disease Control. *Annu.*  
723 *Rev. Phytopathol.* 13, 31–47. <https://doi.org/10.1146/annurev.py.13.090175.000335>

724 Kushalappa, A.C., Chaves, G.M., 1980. An analysis of the development of coffee rust in the field. *Fitopatol.*  
725 *Bras.* 5, 95-103

726 Kushalappa, A.C., Martins, C.P., 1980. Incubation and generation periods for *Hemileia vastatrix* on coffee  
727 in Viçosa, Minas Gerais. *Fitopatol. Bras.* 5, 177-183

728 Kushalappa, A.C., 1981. Linear Models Applied to Variation in the Rate of Coffee Rust Development. *J.*  
729 *Phytopathol.* 101, 22–30. <https://doi.org/10.1111/j.1439-0434.1981.tb03317.x>

730 Kushalappa, A.C., Akutsu, M., Ludwig, A., 1983. Application of Survival Ratio for Monocyclic Process  
731 of *Hemileia vastatrix* in Predicting Coffee Rust Infection Rates. *Phytopathology* 73, 96–103.  
732 <https://doi.org/10.1094/Phyto-73-96>

733 Kushalappa, A.C., Akutsu, M., Oseguera, S.H., Chaves, G.M., Melles, C.A., Miranda, J.M., Bartolo, G.F.,  
734 1984. Equations for predicting the rate of coffee rust development based on net survival ratio for  
735 macrocyclic process of *Hemileia vastatrix*. *Fitopatol. Bras.* 9, 255-271.

736 Kushalappa, A.C., Hernández, T.A., Lemos, H.G., 1986. Evaluation of simple and complex coffee rust  
737 forecasts to time fungicide application. *Fitopatol. Bras.* 11, 515-526.

738 Kushalappa, A.C., Eskes, A.B., 1989. Advances in Coffee Rust Research. *Annu. Rev. Phytopathol.* 27,  
739 503–531.

740 Lin, B.B., 2007. Agroforestry management as an adaptive strategy against potential microclimate  
741 extremes in coffee agriculture. *Agric. Forest Meteorol.* 144, 85–94.

742 Leandro-Muñoz, M.E., Tixier, P., Germon, A., Rakotobe, V., Phillips-Mora, W., Maximova, S., Avelino,  
743 J., 2017. Effects of microclimatic variables on the symptoms and signs onset of *Moniliophthora*  
744 *roreri*, causal agent of Moniliophthora pod rot in cacao. *PLOS ONE* 12, e0184638.  
745 <https://doi.org/10.1371/journal.pone.0184638>

746 Leguizamón-Caycedo, J., Orozco-Gallego, L., Gómez-Gómez, L., 1998. Períodos de incubación (PI) y de  
747 latencia (PL) de la roya del café en la zona central cafetera de Colombia. *Cenicafé* 49, 325–339.

748 Liebig, T., Ribeyre, F., Läderach, P., Poehling, H.-M., van Asten, P., Avelino, J., 2019. Interactive effects  
749 of altitude, microclimate and shading system on coffee leaf rust. *J. Plant Interact.* 14, 407–415.  
750 <https://doi.org/10.1080/17429145.2019.1643934>

751 López-Bravo, D.F., Virginio-Filho, E. de M., Avelino, J., 2012. Shade is conducive to coffee rust as  
752 compared to full sun exposure under standardized fruit load conditions. *Crop Prot.* 38, 21–29.  
753 <https://doi.org/10.1016/j.cropro.2012.03.011>

754 Luaces, O., Rodrigues, L.H.A., Meira, C.A.A., Quevedo, J.R., Bahamonde, A., 2010. Viability of an  
755 alarm predictor for coffee rust disease using interval regression, in: *International Conference on*  
756 *Industrial, Engineering and Other Applications of Applied Intelligent Systems*. Springer, pp. 337–  
757 346.

758 McCain, J.W., Hennen, J.F., 1984. Development of the uredinial thallus and sorus in the orange rust  
759 fungus, *Hemileia vastatrix*. *Phytopathology* 74, 714–721.

760 Meira, C.A., Rodrigues, L.H., Moraes, S.A., 2008. Análise Da Epidemia Da Ferrugem Do Cafeeiro Com  
761 árvore De Decisão. *Trop. Plant Pathol.* 33, 114–124.

762 Meira, C.A.A., Rodrigues, L.H.A., Moraes, S.A. de, 2009. Modelos de alerta para o controle da ferrugem-  
763 do-cafeeiro em lavouras com alta carga pendente. *Pesq. agropec. bras.* 44, 233–242.

764 Merle, I., Pico, J., Granados, E., Boudrot, A., Tixier, P., Virginio Filho, E.M., Cilas, C., Avelino, J., 2019.  
765 Unraveling the complexity of coffee leaf rust behavior and development in different *Coffea*  
766 *arabica* agro-ecosystems. *Phytopathology*. In press.

767 Miller, P.C., 1971. Sampling to Estimate Mean Leaf Temperatures and Transpiration Rates in Vegetation  
768 Canopies. *Ecology* 52, 885–889. <https://doi.org/10.2307/1936038>

769 Nutman, F.J., Roberts, F.M., Bock, K.R., 1960. Method of uredospore dispersal of the coffee leaf-rust  
770 fungus, *Hemileia vastatrix*. *Trans. Brit. Mycol. Soc.* 43, 509–IN6. [https://doi.org/10.1016/S0007-1536\(60\)80033-2](https://doi.org/10.1016/S0007-1536(60)80033-2)

771

772 Nutman, F.J., Roberts, F.M., Clarke, R.T., 1963. Studies on the biology of *Hemileia vastatrix* Berk. & Br.  
773 *Trans. Brit. Mycol. Soc.* 46, 27–44. [https://doi.org/10.1016/S0007-1536\(63\)80005-4](https://doi.org/10.1016/S0007-1536(63)80005-4)

774 Pedro, M.J., 1983. Effects of Meteorological Factors on the Development of Coffee Leaf Rust. *EPPO*  
775 *Bull.* 13, 153–155. <https://doi.org/10.1111/j.1365-2338.1983.tb01592.x>

776 Perez-Ariza, C.B., Nicholson, A.E., Flores, M.J., 2012. Prediction of coffee rust disease using bayesian  
777 networks, in: *Sixth European Workshop on Probabilistic Graphical Models*. Granada (Spain), pp.  
778 259–266.

779 Pezzopane, J.R.M., Souza, P.S. de, Rolim, G.D.S., Gallo, P.B., 2011. Microclimate in coffee plantation  
780 grown under grevillea trees shading. *Acta Sci. Agron.* 33, 201–206.  
781 <https://doi.org/10.4025/actasciagron.v33i2.7065>

- Pietravalle, S., Shaw, M.W., Parker, S.R., van den Bosch, F., 2003. Modeling of relationships between weather and *Septoria tritici* epidemics on winter wheat: A critical approach. *Phytopathology* 93,1329-1339.
- Pinto, A.C.S., Pozza, E.A., de Souza, P.E., Pozza, A.A.A., Talamini, V., Boldini, J.M., Santos, F.S., 2002. Description of epidemics of coffee rust with neural networks. *Fitopatol. Bras.* 27, 517–524. <https://doi.org/10.1590/S0100-41582002000500013>
- Rayner, R.W., 1961. Germination and penetration studies on coffee rust (*Hemileia vastatrix* B. & Br.). *Annals of Applied Biology* 49, 497–505. <https://doi.org/10.1111/j.1744-7348.1961.tb03641.x>
- Ribeiro, I.J.A., 1978. Efeito de alta temperatura no desenvolvimento de *Hemileia vastatrix* em cafeeiro suscetível. *Bragantia* 37, 11–16.
- Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., Müller, M., 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 12, 77. <https://doi.org/10.1186/1471-2105-12-77>
- Rudin, C., 2019. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence* 1, 5, 206.
- Sache, I., 2000. Short-distance dispersal of wheat rust spores by wind and rain. *Agronomie* 20, 757–767.
- Santacreo, R., Polanco, E., Oseguera, S., 1983. Periodo de incubación y generación de *Hemileia vastatrix* Berk. & Br. en tres zonas cafetaleras de Honduras, Centro América, in: VI Simposio Latinoamericano sobre Caficultura, IICA: Panamá (Panamá). 109-127.
- Savary, S., Janeau, J.L., Alloreant, D., Escalante, M., Avelino, J., Willocquet, L., 2004. Effects of simulated rainfall events on spore dispersal and spore stocks in three tropical pathosystems. *Phytopathology* 94:S92
- Schindelin, J., Rueden, C.T., Hiner, M.C., Eliceiri, K.W., 2015. The ImageJ ecosystem: An open platform for biomedical image analysis: THE IMAGEJ ECOSYSTEM. *Mol. Reprod. Dev.* 82, 518–529. <https://doi.org/10.1002/mrd.22489>
- Siettos, C.I., Russo, L., 2013. Mathematical modeling of infectious disease dynamics. *Virulence* 4, 295–306. <https://doi.org/10.4161/viru.24041>
- Toniutti, L., Breitler, J.-C., Etienne, H., Campa, C., Doulbeau, S., Urban, L., Lambot, C., Pinilla, J.-C.H., Bertrand, B., 2017. Influence of Environmental Conditions and Genetic Background of Arabica Coffee (*C. arabica* L) on Leaf Rust (*Hemileia vastatrix*) Pathogenesis. *Front. Plant Sci.* 8. <https://doi.org/10.3389/fpls.2017.02025>
- Tronconi, N., Palma, M.R., Suazo, G., Zaldívar, R., Agurcia, R.D., 1995. Periodo de incubación y generación de *Hemileia vastatrix* Berk & Br. en Honduras, in: V Seminario Nacional de Investigación y de Transferencia en Caficultura, IHCAFE: Tegucigalpa (Honduras). 165-173.
- Villegas-García, C., Baeza-Aragón, C.A., 1990. Diseminación de *Hemileia vastatrix* Berk. y Br. a nivel del árbol, en un foco natural. *Cenicafé* 41, 39–49.
- Waller, J.M., 1982. Coffee rust-epidemiology and control. *Crop Prot.* 1, 385–404. [https://doi.org/10.1016/0261-2194\(82\)90022-9](https://doi.org/10.1016/0261-2194(82)90022-9)

## Titles for figures



**Fig. 1.** Symptoms and signs of *Hemileia vastatrix* on the under-side of a coffee leaf. Photographs by Steven Cerdas Hernandez taken in Turrialba the 7th, the 14th and 28th of June 2017 from left to right.

**Fig. 2.** Graphics showing daily rainfall (black bar plot) and daily minimum (blue points) and maximum (orange points) air temperatures measured by the weather stations in the four plots from May 2017 to July 2018. In Turrialba, for both coffee plots in full sun and under medium shade, there was only one rain gauge, set in the full sun coffee plot.

**Fig. 3.** Inoculum stock influence on lesion emergence probability, per periods from one to eight weeks before the symptom detection (abscises) and for durations from one to eight weeks (ordinates). The scale on the right indicates the absolute difference between Akaike information criterion (AIC) of the model including the inoculum stock averaged over the considered periods (dates and durations) ( $NewLesion \sim LeafStratum + Inoculum_{period}$ ) and AIC of its reference model without including the inoculum stock ( $NewLesion \sim LeafStratum$ ). The highest AIC difference, labelled with a white square, indicates the period to be considered for inoculum stock that better explains the lesion emergence probability: 5 weeks before lesion emergence for 3 weeks. Code explanations are available in Table 2.

**Fig. 4.** Influence of microclimatic variables on lesion emergence probability by foliar site, per periods from one to 40 days before the symptom detection (abscises) and for durations from one to 40 days (ordinates). The scale on the right indicates the absolute difference between Akaike information criterion (AIC) of the model including the microclimatic variable averaged over the considered periods (dates and durations) ( $NewLesions \sim PastInoculum + MicroclimVar_{period}$ ) and AIC of its reference model without including the microclimatic variable ( $NewLesions \sim PastInoculum$ ). The highest AIC differences, labelled with black circles, indicate the periods, to be considered for the microclimatic variable, that better explain the lesion emergence probability.

**Fig. 5.** Influence of microclimatic variables on sporulation probability, per periods from one to 28 days before the sign detection (abscises) and for durations from one to 28 days (ordinates). The scale on the right indicates the absolute difference between Akaike information criterion (AIC) of the model including the microclimatic variable averaged over the considered periods (dates and durations) ( $Sporulation \sim LesionAge + MicroclimVar_{period}$ )

and AIC of its reference model without including the microclimatic variable (*Sporulation* ~ *LesionAge*). The highest AIC differences, labelled with black circles, indicate the periods, to be considered for the microclimatic variable, that better explain sporulation probability.

**Fig. 6.** Influence of microclimatic variables on lesion infectious area growth, per periods from one to seven days before the lesion infectious area growth (abscises) and for durations from one to seven days (ordinates). The scale on the right indicates the absolute difference between Akaike information criterion (AIC) of the model including the microclimatic variable averaged over the considered periods (dates and durations) (*InfectiousArea* ~ *InfectiousArea*<sub>day-7</sub> + *MicroclimVar*<sub>period</sub>) and AIC of its reference model without including the microclimatic variable (*InfectiousArea* ~ *InfectiousArea*<sub>day-7</sub>). The highest AIC differences, labelled with black circles, indicate the periods, to be considered for the microclimatic variable, that better explain lesion infectious area growth.

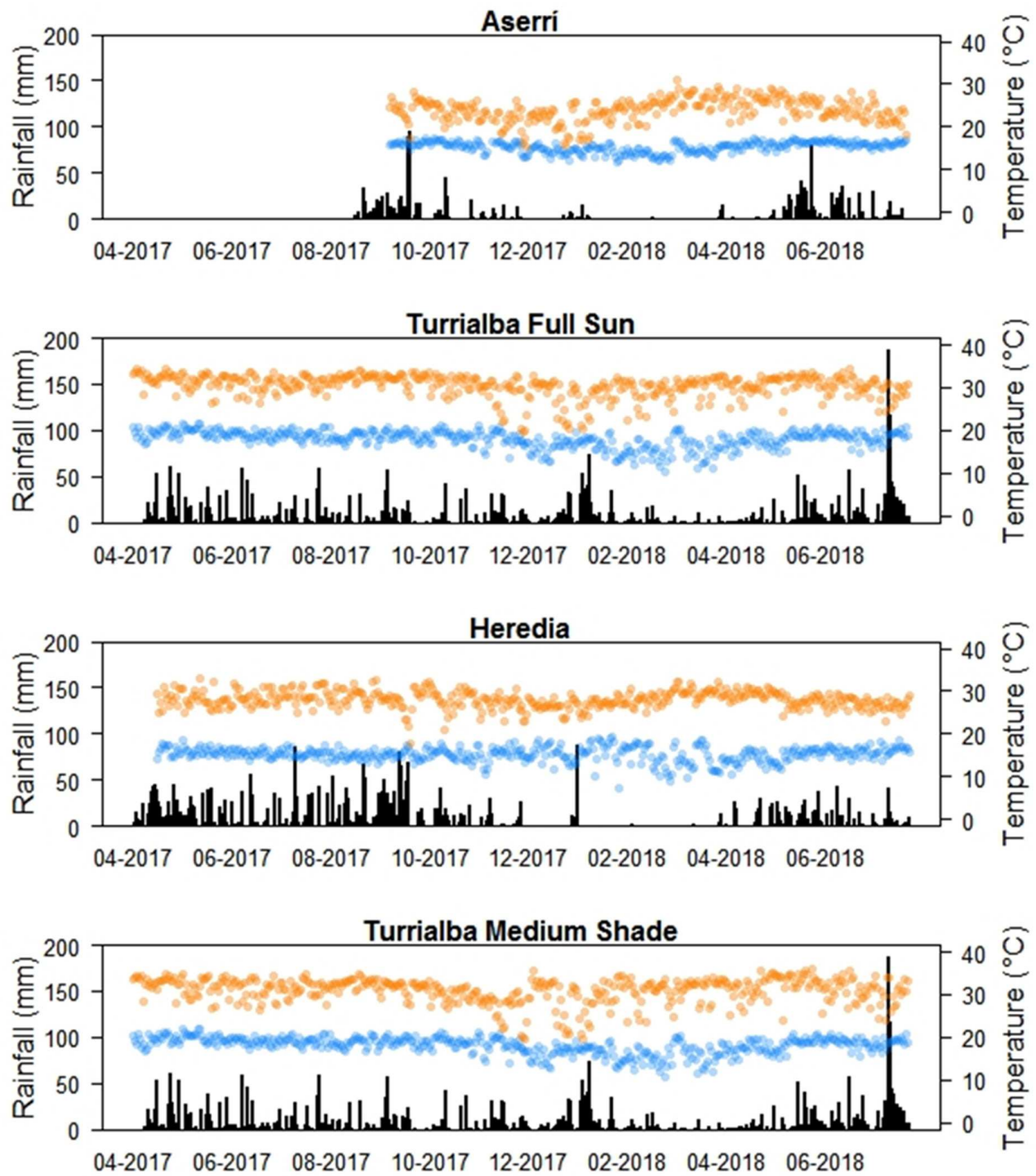
**Fig. 7.** Graphs displaying the distributions of the microclimatic variables used to develop the complete model forecasting the lesion emergence probability by foliar site. Grey areas are removed ranges due to under representation of data. *TcMin*<sub>T20D3</sub> : averaged daily minimum leaf temperature over a period of three days starting 20 days before lesion emergence; *Rain*<sub>T33D10</sub> : averaged daily rainfall over a period of 10 days starting 33 days before lesion emergence; *TcAmp*<sub>T20D11</sub> : averaged daily amplitude of leaf temperature over a period of 11 days starting 20 days before lesion emergence.

**Fig. 8.** Predictions of models A1 and A2 (graphs A1 and A2) for the lesion emergence probability of one of the 25 leaf sites considered, model B for sporulation probability (graph B) and model C for lesion infectious area growth (graph C). Graphs A1 and A2 were presented among model's variation ranges of *Rain*<sub>T33D10</sub> and *TcMin*<sub>T20D3</sub>, while *TcAmp*<sub>T20D11</sub> (ranges [6.2:17.9] in model A1 and [7.1:17.3] in model A2) was fixed to its mean and *PastInoculum* (range [0:1.1] in model A1 and [0:3.8] in model A2) to its first (clear grey) and third quantiles (dark grey). Graph B was represented among variation ranges of *TaMin*<sub>T15D4</sub> and *Rain*<sub>T5D3</sub>, while *Rain*<sub>T12D2</sub> (range [0:28.3]), *TcAmp*<sub>T4D1</sub> (range [5.9:18.9]) and *PlantFruitLoad* (range [0:1150]) were fixed to their mean and *LesionAge* (range [0:50]) was fixed to seven (clear grey), fifteen (grey) and 23 days (dark grey). Graph C was presented among variation ranges of *TaMax*<sub>T5D1</sub> and *TcAmp*<sub>T1D1</sub> and *InfectiousArea*<sub>day-7</sub>

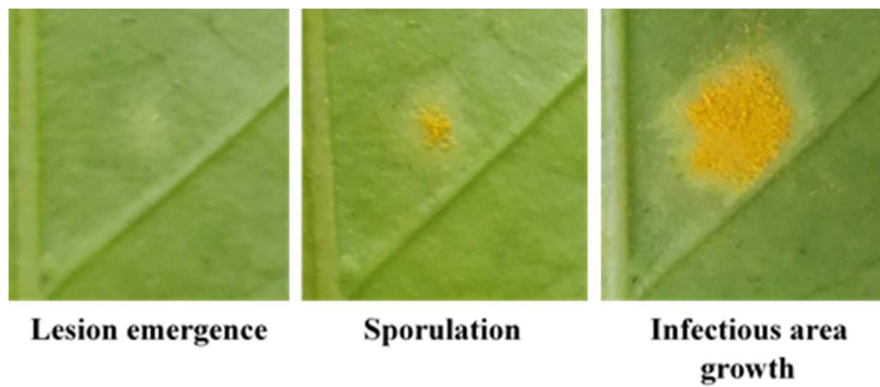
(range [0:0.295]) was fixed to its first (clear grey) and third quantiles (dark grey). All the variables are described in Tables 2 and 4.

**Fig. 9.** Regressions between predicted values and observed averaged values by plot and date for new lesion emergence probability by foliar site (A), sporulation probability (B) and lesion infectious area (C).

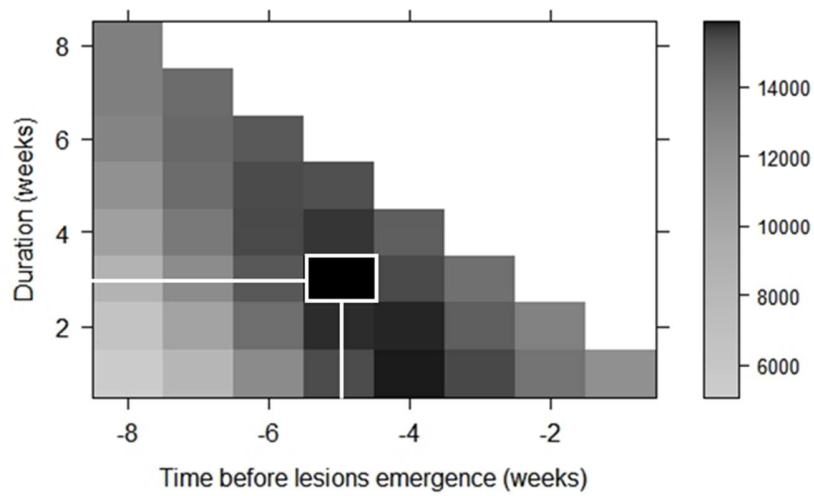
**Fig. 10.** Representations with a time axis of the variables corresponding to the complete models A1, A2, B and C. The box color of the variables indicates a positive effect (white), a negative effect (black) and a unimodal effect with a maximum (grey).



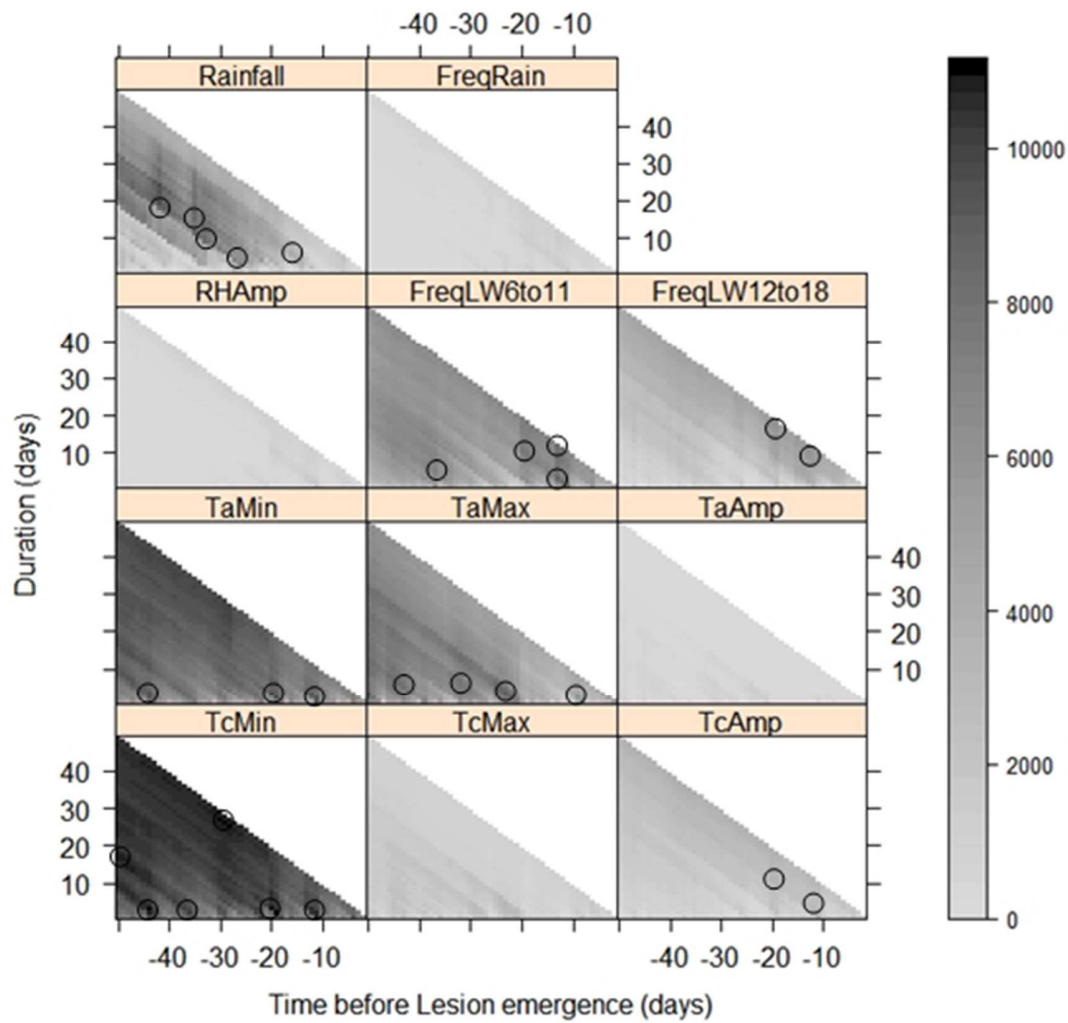
**Fig. 1.** Graphics showing daily rainfall (black bar plot) and daily minimum (blue points) and maximum (orange points) air temperatures measured by the weather stations in the four plots from May 2017 to July 2018. In Turrialba, for both coffee plots in full sun and under medium shade, there was only one rain gauge, set in the full sun coffee plot.



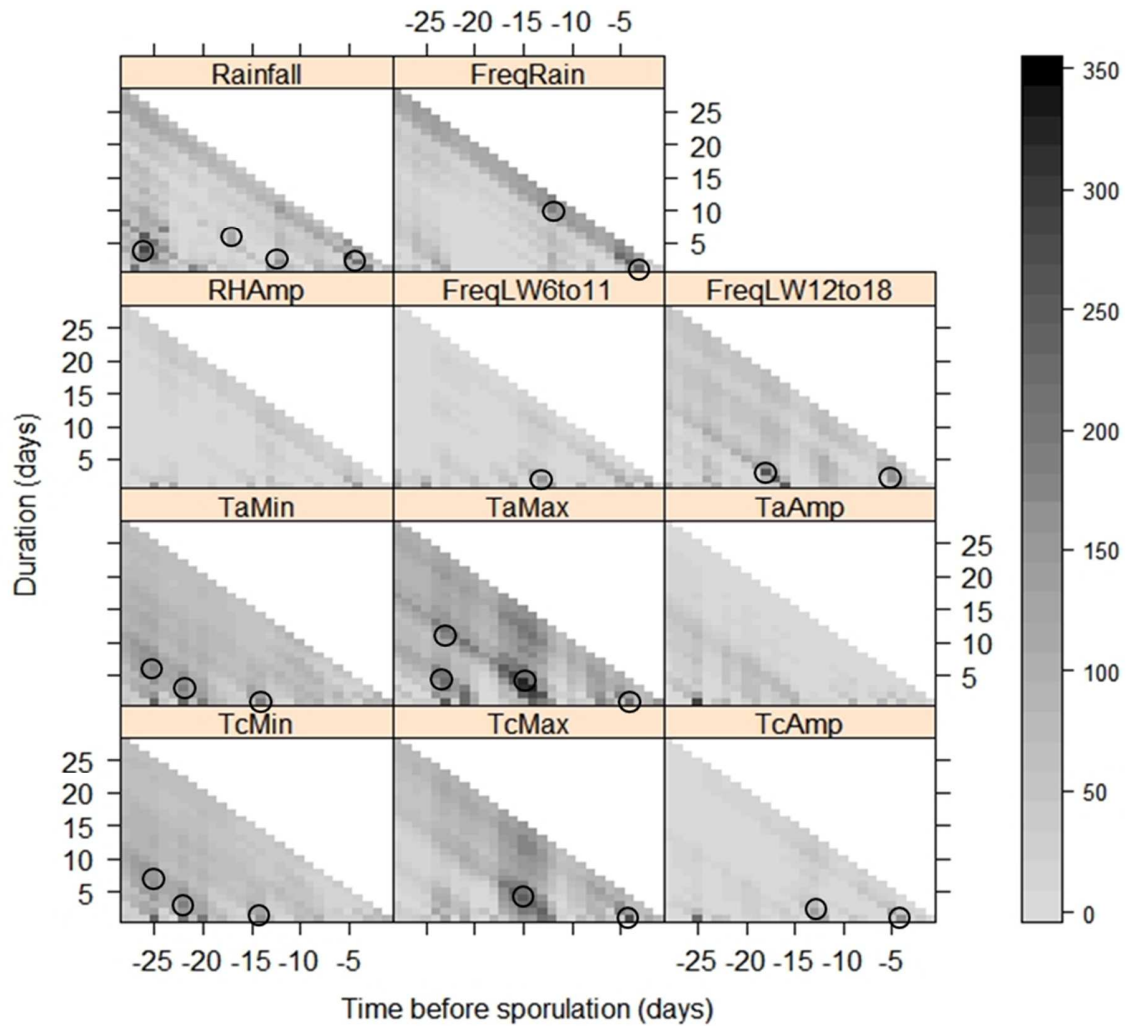
**Fig. 2.** Symptoms and signs of *Hemileia vastatrix* on the under-side of a coffee leaf. Photographs by Steven Cerdas Hernandez taken in Turrialba the 7th, the 14th and 28th of June 2017 from left to right.



**Fig. 3.** Inoculum stock influence on lesion emergence probability, per periods from one to eight weeks before the symptom detection (abscises) and for durations from one to eight weeks (ordinates). The scale on the right indicates the absolute difference between Akaike information criterion (AIC) of the model including the inoculum stock averaged over the considered periods (dates and durations) ( $NewLesion \sim LeafStratum + Inoculum_{period}$ ) and AIC of its reference model without including the inoculum stock ( $NewLesion \sim LeafStratum$ ). The highest AIC difference, labelled with a white square, indicates the period to be considered for inoculum stock that better explains the lesion emergence probability: 5 weeks before lesion emergence for 3 weeks. Code explanations are available in Table 2.

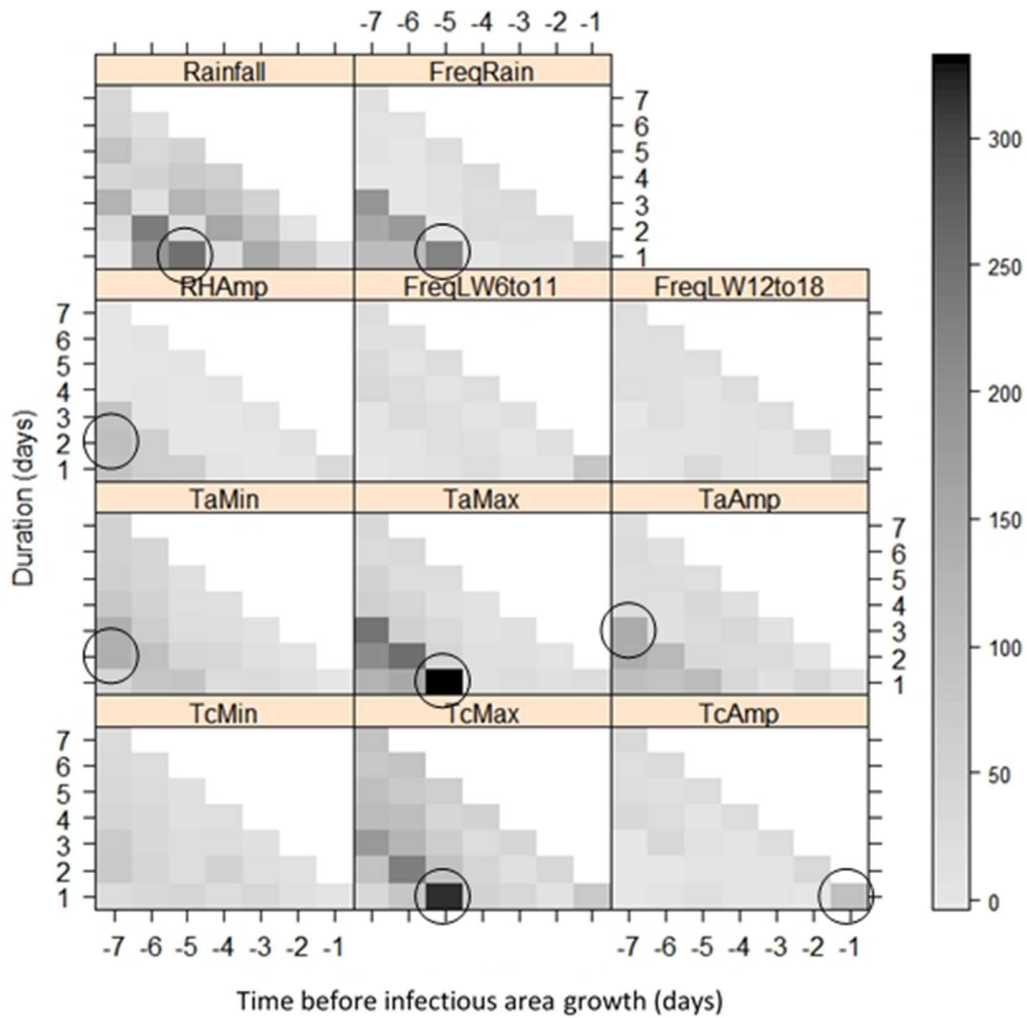


**Fig. 4.** Influence of microclimatic variables on lesion emergence probability by foliar site, per periods from one to 50 days before the symptom detection (abscises) and for durations from one to 50 days (ordinates). The scale on the right indicates the absolute difference between Akaike information criterion (AIC) of the model including the microclimatic variable averaged over the considered periods (dates and durations) ( $NewLesions \sim PastInoculum + MicroclimVar_{period}$ ) and AIC of its reference model without including the microclimatic variable ( $NewLesions \sim PastInoculum$ ). The highest AIC differences, labelled with black circles, indicate the periods, to be considered for the microclimatic variable, that better explain the lesion emergence probability.

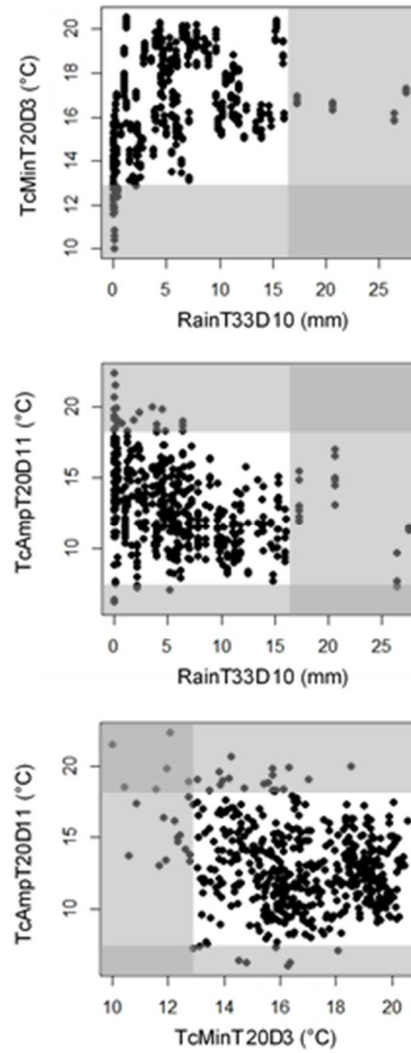


**Fig. 5.** Influence of microclimatic variables on sporulation probability, per periods from one to 28 days before the sign detection (abscises) and for durations from one to 28 days (ordinates). The scale on the right indicates the absolute difference between Akaike information criterion (AIC) of the model including the microclimatic variable averaged over the considered periods (dates and durations) ( $Sporulation \sim LesionAge + MicroclimVar_{period}$ ) and AIC of its reference model without including the microclimatic variable ( $Sporulation \sim LesionAge$ ). The highest AIC differences, labelled with black circles, indicate the periods, to be considered for the microclimatic variable, that better explain sporulation probability.

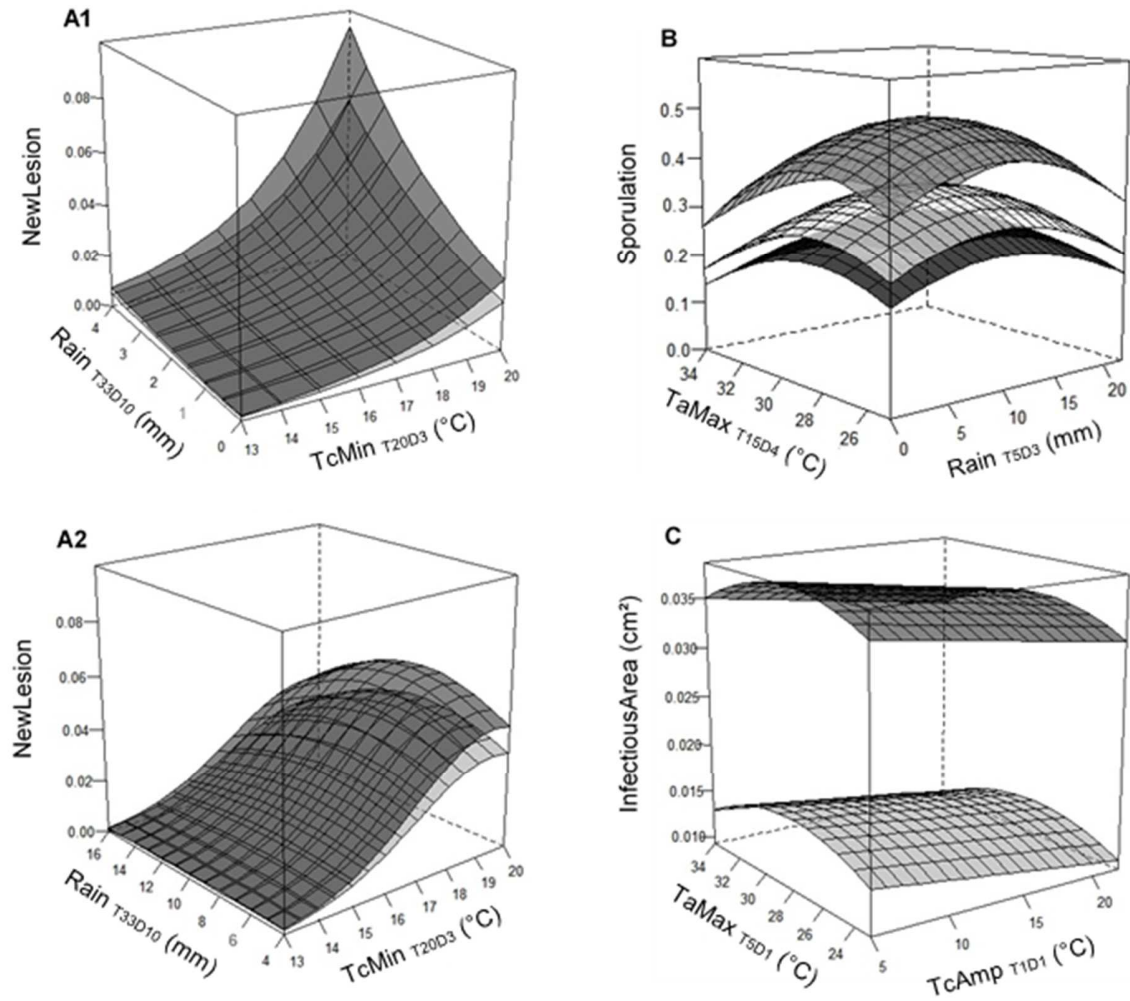




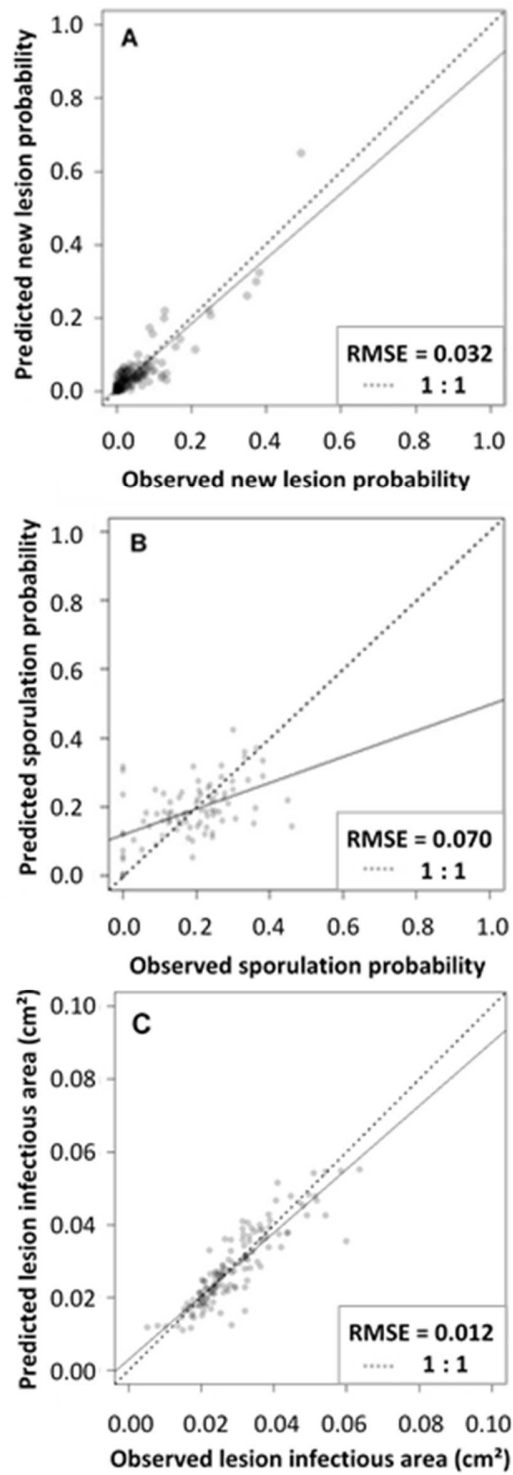
**Fig. 6.** Influence of microclimatic variables on lesion infectious area growth, per periods from one to seven days before the lesion infectious area growth (abscises) and for durations from one to seven days (ordinates). The scale on the right indicates the absolute difference between Akaike information criterion (AIC) of the model including the microclimatic variable averaged over the considered periods (dates and durations) ( $InfectiousArea \sim InfectiousArea_{day-7} + MicroclimVar_{period}$ ) and AIC of its reference model without including the microclimatic variable ( $InfectiousArea \sim InfectiousArea_{day-7}$ ). The highest AIC differences, labelled with black circles, indicate the periods, to be considered for the microclimatic variable, that better explain lesion infectious area growth.



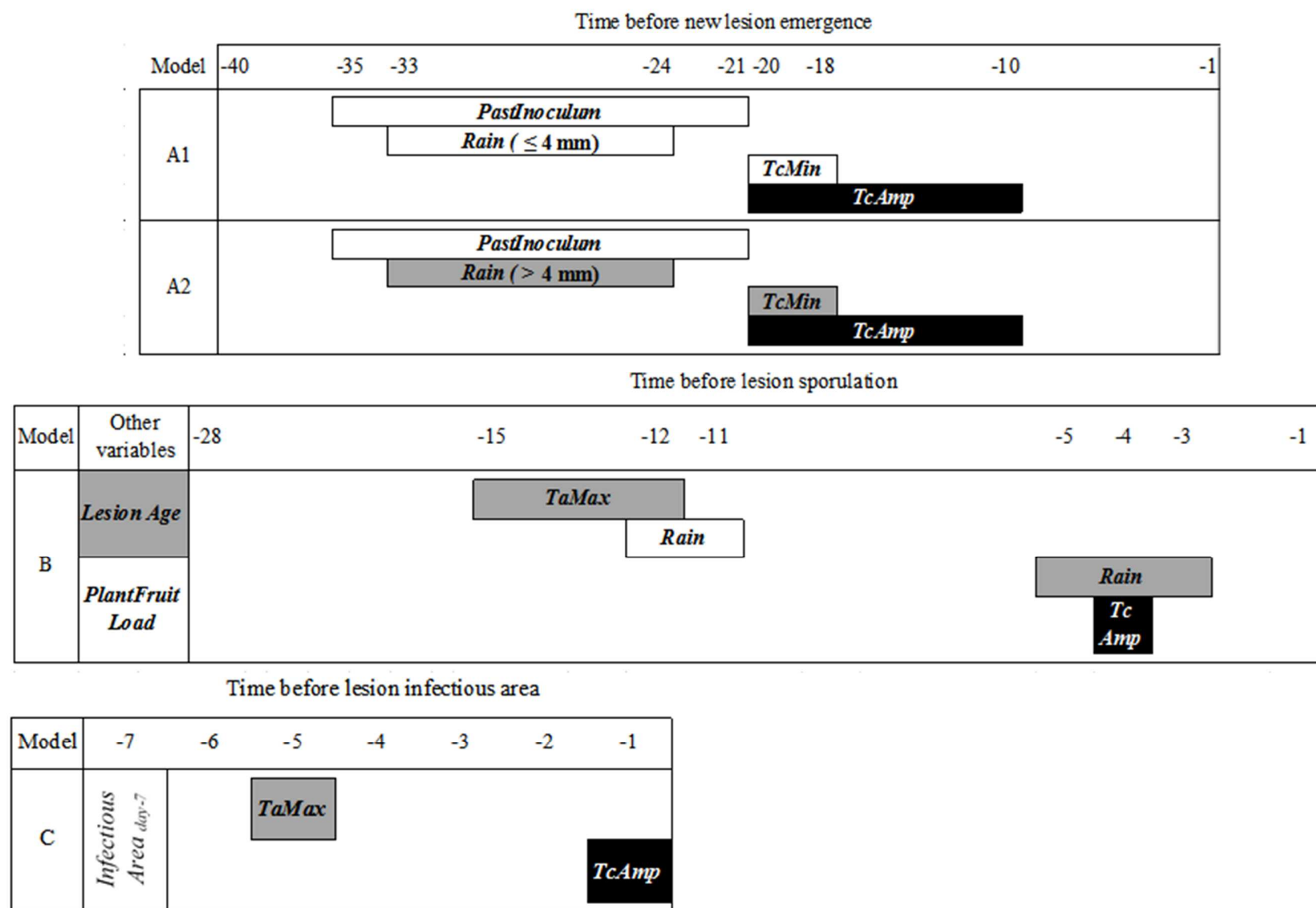
**Fig. 7.** Graphs displaying the distributions of the microclimatic variables used to develop the complete model forecasting the lesion emergence probability by foliar site. Grey areas are removed ranges due to under representation of data.  $TcMinT20D3$  : averaged daily minimum leaf temperature over a period of three days starting 20 days before lesion emergence;  $RainT33D10$  : averaged daily rainfall over a period of 10 days starting 33 days before lesion emergence;  $TcAmpT20D11$  : averaged daily amplitude of leaf temperature over a period of 11 days starting 20 days before lesion emergence.



**Fig. 8.** Predictions of models A1 and A2 (graphs A1 and A2) for the lesion emergence probability of one of the 25 leaf sites considered, model B for sporulation probability (graph B) and model C for lesion infectious area growth (graph C). Graphs A1 and A2 were presented among model's variation ranges of  $Rain_{T33D10}$  and  $TcMin_{T20D3}$ , while  $TcAmp_{T20D11}$  (ranges [6.2:17.9] in model A1 and [7.1:17.3] in model A2) was fixed to its mean and  $PastInoculum$  (range [0:1.1] in model A1 and [0:3.8] in model A2) to its first (clear grey) and third quantiles (dark grey). Graph B was represented among variation ranges of  $TaMin_{T15D4}$  and  $Rain_{T5D3}$ , while  $Rain_{T12D2}$  (range [0:28.3]),  $TcAmp_{T4D1}$  (range [5.9:18.9]) and  $PlantFruitLoad$  (range [0:1150]) were fixed to their mean and  $LesionAge$  (range [0:50]) was fixed to seven (clear grey), fifteen (grey) and 23 days (dark grey). Graph C was presented among variation ranges of  $TaMax_{T5D1}$  and  $TcAmp_{T1D1}$  and  $InfectiousArea_{day-7}$  (range [0:0.295]) was fixed to its first (clear grey) and third quantiles (dark grey). All the variables are described in Tables 2 and 3.



**Fig. 9.** Regressions between predicted values and observed averaged values by plot and date for new lesion emergence probability by foliar site (A), sporulation probability (B) and lesion infectious area (C).



**Fig. 10.** Representations with a time axis of the variables corresponding to the complete models A1, A2, B and C. The box color of the variables indicates a positive effect (white), a negative effect (black) and a unimodal effect with a maximum (grey).

**Table 1**

Weather-based predictive models of coffee leaf rust development

Variable to forecast	Studied period for climatic variables	Number of tested variables			References
		Weather	Rust	Host and practices	
Latency period	Latency period	2	0	0	Rayner 1961 Kushalappa and Martins 1980 Santacreo et al. 1983 Tronconi et al. 1995
		2	0	1	de Moraes et al. 1976
Incubation period	Incubation period	2	0	0	Kushalappa and Martins 1980 Tronconi et al. 1995
Number of lesions per leaf	Over last 15, 30 or 45 days	3	0	1	Alfonsi et al. 1974
Proportion of rust area	Over the last 40 days	1	2	0	Kushalappa and Chaves 1980
Apparent infection rate over 28 days	Over last 14 days or from 28 to 14 days before	4	2	1	Kushalappa 1981
	Semi-mechanistic approach	4	2	1	Kushalappa et al. 1983
		4	2	2	Kushalappa et al. 1984 Holguín 1987
Apparent infection rate over 15 days	Over last 15 days	5	0	0	Pedro 1983
Incidence	Over last 15, 30, 45 or 60 days	8	0	1	Pinto et al. 2002
	Over the year	1	1	> 40	Avelino et al. 2006
Class of monthly incidence variations	Latency period	> 10	0	2	Meira et al. 2008
		> 10	0	1	Meira et al. 2009
		> 10	1	1	Cintra et al. 2011
Class of incidence	Over last 45 days	> 10	1	2	Luaces et al. 2010
		> 20	1	2	Perez Ariza et al. 2012
	Last month and last two months	6	0	7	Corrales et al. 2015, 2016
Month average of daily infection risk	Mechanistic approach	2	0	0	Bebber et al. 2016
Apparent infection rate over 30 days	Periods of 30 days, 90 days before, 60 days before and 30 days before	> 10	0	2	Hinnah et al. 2018
Incidence	Last 28 days	4	0	6	Liebig et al., 2019

**Table 2**

Description of the three forecasted variables on rust development and their explanatory variables including microclimatic descriptors

Dependent variables	Description	Unit	Range
NewLesion	Presence/absence of a new rust lesion per leaf site	-	0 or 1
Sporulation	Presence/absence of uredospores on each rust lesion	-	0 or 1
InfectiousArea	Area with uredospores per lesion	cm <sup>2</sup>	[0.001;0.535]
Explanatory variables	Description	Unit	Range
Inoculum	Area with uredospores per branch averaged per plot	cm <sup>2</sup>	[0;3.9]
LesionAge	Time since the lesion emergence	days	0, 7, 14, 21, 28, 35, 42, 49
InfectiousArea <sub>day-7</sub>	Past area with uredospores per lesion (7 days ago)	cm <sup>2</sup>	[0.001;0.295]
LeafStratum	Coffee leaf stratum (Bottom; Middle; Top)	-	-
PlantFruitLoad	Plant fruiting nodes	-	[0;1793]
Rainfall	Daily rainfall	mm	[0;186.9]
FreqRain	Daily number of rainy hours	-	[0;24]
TaMin	Daily minimum air temperature	°C	[7.5;21.9]
TaMax	Daily maximum air temperature	°C	[15.5;35.9]
TaAmp	Daily amplitude of air temperature	°C	[1.9;22.9]
RHAmp	Daily amplitude of relative humidity	%	[0;87.8]
TcMin	Daily minimum leaf temperature	°C	[8.4;22.0]
TcMax	Daily maximum leaf temperature	°C	[15.5;41.3]
TcAmp	Daily amplitude of leaf temperature	°C	[1.6;28.8]
FreqLW6to11	Leaf wetness duration from 6am to 11am	h	[0;6]
FreqLW12to18	Leaf wetness duration from 12am to 6pm	h	[0;7]

**Table 3**

Microclimatic variables identified as the most explicative for each of the three models of rust development. The variables in bold are the ones conserved after removing the over-correlated variables ( $r > 0.7$ ) and those underlined are the variables used in the full models.

Model	Microclimatic variables identified			
Lesion emergence probability	TaMin <sub>T11D1</sub>	TcMin <sub>T11D1</sub>	Rain <sub>T16D6</sub>	FreqLW6to11 <sub>T13D3</sub>
	TaMin <sub>T19D2</sub>	<b><u>TcMin<sub>T20D3</sub></u></b>	Rain <sub>T26D3</sub>	FreqLW6to11 <sub>T13D12</sub>
	TaMin <sub>T44D3</sub>	TcMin <sub>T28D25</sub>	<b><u>Rain<sub>T33D10</sub></u></b>	FreqLW6to11 <sub>T20D11</sub>
		TcMin <sub>T37D3</sub>	Rain <sub>T35D17</sub>	FreqLW6to11 <sub>T37D5</sub>
	TaMax <sub>T9D2</sub>	TcMin <sub>T44D3</sub>	Rain <sub>T42D19</sub>	
	TaMax <sub>T23D3</sub>	TcMin <sub>T50D17</sub>		FreqLW12to18 <sub>T13D10</sub>
	TaMax <sub>T32D6</sub>			FreqLW12to18 <sub>T19D16</sub>
	TaMax <sub>T44D6</sub>	TcAmp <sub>T11D2</sub> <b><u>TcAmp<sub>T20D11</sub></u></b>		
Sporulation probability	TaMin <sub>T14D1</sub>	TcMin <sub>T14D1</sub>	<b><u>Rain<sub>T5D3</sub></u></b>	FreqLW6to11 <sub>T13D2</sub>
	TaMin <sub>T22D3</sub>	TcMin <sub>T22D3</sub>	<b><u>Rain<sub>T12D2</sub></u></b>	
	TaMin <sub>T25D7</sub>	TcMin <sub>T25D6</sub>	<b><u>Rain<sub>T17D5</sub></u></b>	FreqLW12to18 <sub>T5D2</sub>
			Rain <sub>T26D4</sub>	<b>FreqLW12to18<sub>T18D3</sub></b>
	TaMax <sub>T4D1</sub>	TcMax <sub>T4D1</sub>		
	<b><u>TaMax<sub>T15D4</sub></u></b>	TcMax <sub>T15D4</sub>		FreqRain <sub>T3D1</sub>
	TaMax <sub>T23D4</sub>			FreqRain <sub>T12D10</sub>
	TaMax <sub>T23D12</sub>	<b><u>TcAmp<sub>T4D1</sub></u></b> <b><u>TcAmp<sub>T13D2</sub></u></b>		
Lesion infectious area	<b>TaMin<sub>T7D2</sub></b>	TaAmp <sub>T7D3</sub>	<b><u>TcAmp<sub>T1D1</sub></u></b>	Rain <sub>T5D1</sub>
	<b><u>TaMax<sub>T5D1</sub></u></b>	TcMax <sub>T5D1</sub>	RHAmp <sub>T7D2</sub>	FreqRain <sub>T5D1</sub>

Microclimatic variables (MV): TaMin; TaMax; TaAmp; TcMin; TcMax; TcAmp; RHAmp; FreqLW6to11; FreqLW12to18 are described in Table 2 and Rain is a shorter term for Rainfall

MV<sub>TxDy</sub>: averaged daily microclimatic variable over a y-days period starting x days before the symptom or sign onset



**Table 4**

Description of the parameters estimate of the models A1,  
A2, B and C

Model	Parameter	Value [ $\pm$ Standard error]	Unit
A1	$\alpha_{1.1}$	- 11.41 [ $\pm 0.22$ ]	-
	$\alpha_{1.2}$	+ 1.84 [ $\pm 0.065$ ]	cm <sup>-2</sup>
	$\alpha_{1.3}$	+ 0.33 [ $\pm 0.015$ ]	mm <sup>-1</sup>
	$\alpha_{1.4}$	+ 0.38 [ $\pm 0.0099$ ]	°C <sup>-1</sup>
	$\alpha_{1.5}$	- 0.017 [ $\pm 0.0089$ ]	°C <sup>-1</sup>
A2	$\alpha_{2.1}$	- 32.57 [ $\pm 1.22$ ]	-
	$\alpha_{2.2}$	+ 1.15 [ $\pm 0.010$ ]	cm <sup>-2</sup>
	$\alpha_{2.3}$	+ 0.16 [ $\pm 0.015$ ]	mm <sup>-1</sup>
	$\alpha_{2.4}$	- 0.0091 [ $\pm 0.00077$ ]	mm <sup>-2</sup>
	$\alpha_{2.5}$	+ 3.08 [ $\pm 0.14$ ]	°C <sup>-1</sup>
	$\alpha_{2.6}$	- 0.079 [ $\pm 0.0039$ ]	°C <sup>-2</sup>
	$\alpha_{2.7}$	- 0.10 [ $\pm 0.0043$ ]	°C <sup>-1</sup>
B	$\beta_1$	- 19.69 [ $\pm 2.11$ ]	-
	$\beta_2$	+ 0.30 [ $\pm 0.0047$ ]	days <sup>-1</sup>
	$\beta_3$	- 0.010 [ $\pm 0.00021$ ]	days <sup>-2</sup>
	$\beta_4$	+ 0.00016 [ $\pm 0.000044$ ]	-
	$\beta_5$	+ 1.24 [ $\pm 0.14$ ]	°C <sup>-1</sup>
	$\beta_6$	- 0.022 [ $\pm 0.0024$ ]	°C <sup>-2</sup>
	$\beta_7$	+ 0.018 [ $\pm 0.0019$ ]	mm <sup>-1</sup>
	$\beta_8$	+ 0.065 [ $\pm 0.0070$ ]	mm <sup>-1</sup>
	$\beta_9$	- 0.0032 [ $\pm 0.00037$ ]	mm <sup>-2</sup>
	$\beta_{10}$	- 0.045 [ $\pm 0.0043$ ]	°C <sup>-1</sup>
C	$\gamma_1$	- 0.068 [ $\pm 0.0081$ ]	-
	$\gamma_2$	+ 1.30 [ $\pm 0.0045$ ]	cm <sup>-2</sup>
	$\gamma_3$	+ 0.0057 [ $\pm 0.00057$ ]	°C <sup>-1</sup>
	$\gamma_4$	- 0.00010 [ $\pm 0.000010$ ]	°C <sup>-2</sup>
	$\gamma_5$	- 0.00021 [ $\pm 0.000030$ ]	°C <sup>-1</sup>