Eur J Plant Pathol (2010) 126:349–362
DOI 10.1007/s10658-009-9540-1

Segmentation applied to weather-disease relationships in South American leaf blight of the rubber tree

Jean Guyot · Virgile Condina · Fabien Doaré · Christian Cilas · Ivan Sache

Received: 15 May 2009 / Accepted: 28 September 2009 / Published online: 16 October 2009

Abstract South American leaf blight (SALB) is a severe threat to world rubber production. One way of controlling it is to set up plantations in zones not conducive to the disease. Such zones are known once a plantation has been set up, but few data are available on how climate affects the disease, especially in the Amazon region. With better knowledge of conditions that are favourable to SALB epidemics it would be possible to more accurately identify risk zones in Asia and Africa, continents that are still SALB-free. Based on a trial design involving detailed and frequent observations, and with a method rarely used in plant epidemiology, the segmentation method, the results presented in this article make it possible to list, in order of importance, climatic factors that influence disease severity under conditions where the climate varies little over the year.

Keywords Amazonia · CART · Epidemiology · French Guiana · *Hevea brasiliensis* · *Microcyclus ulei* · SALB

J. Guyot · V. Condina · F. Doaré
CIRAD, UPR Bioagresseurs de Pérennes,
97 387 Kourou, Cédex, France

C. Cilas
CIRAD, UPR Bioagresseurs de Pérennes,
34 398 Montpellier, Cédex 5, France

I. Sache
INRA—AgroParisTech,
UMR1290 Bioger-CPP, BP 01, 78 850 Thiverval-Grignon,
France

J. Guyot (✉)
CIRAD,
PB 701, 97387 Kourou Cédex, France
e-mail: jean.guyot@cirad.fr

Introduction

South American leaf blight (SALB), caused by the Ascomycete fungus *Microcyclus ulei* (Dothideales, anamorph: *Fusicladium macrosporum* teleomorphs: *Dothidella ulei* or *Melanopsammopsis ulei*), is a major obstacle to rubber growing in South and Central America. Furthermore, the disease would be a severe threat to natural rubber production if it were to be introduced into Asia and Africa (Chee 1985). Conidia and ascospores of the fungus infect immature leaves and can cause their abscission. In susceptible clones, trees defoliated for several consecutive years grow slowly, rubber production is low and trees can eventually die. While there is currently no efficient method of disease control, promising research includes the selection of resistant clones (Garcia et al. 2004) and the definition of escape areas less favourable to SALB, where rubber cultivation could be recommended.

Under controlled conditions, spore germination, growth and sporulation of the fungus are favoured by temperatures between 24°C and 28°C (Holliday 1970; Gasparotto et al. 1989a). Long wet periods enhance

disease development while an excess of free water inhibits spore germination and mycelial growth. (Chee et al. 1985), but these parameters are not quantified. In the field, in locations with contrasting climatic conditions, a temperature <20°C has a negative effect on disease development, while high relative humidity (RH) (number of days with wetness duration >90% during 6 h consecutively) and longer wet periods (number of days with wetness duration >6 h consecutively) are highly correlated with diseased leaf area. Rainfall has been shown to have no significant effect on disease development (De Camargo et al. 1967; Gasparotto 1988; Gasparotto et al. 1989b) but some authors observe less damage during periods of heavy rainfall but they do not give any quantitative information for this detrimental effect to disease development (Stahel 1917; Chee 1976b; Chee 1980; Holliday 1969). In Amazonia, from where the genus *Hevea* and its pathogen *M. ulei* both originate, the weather conditions are known to be conducive to SALB all year round (Gasparotto et al. 1989b). However there are no data on how they affect variations in disease severity (Gasparotto et al. 1991). Moreover, the few available studies on weather effects on disease development hardly mention any interactions between weather variables. Gasparotto (1988) observed that variations in disease level were not always explained by the expected effects of individual weather variables, but did not investigate further the origin of those variations. Gasparotto observed in Viana (Esperito Santo, Brazil) that 1. the damage was severe during a season with long periods with RH >90% but with temperatures almost permanently <20°C, sometimes near 10°C, with an average between 17°C and 25°C, and maximum temperatures rarely >30°C. 2. in another period of the year, with RH identical to the previous ones, but with higher temperatures (average 25–30°C, minimum very rarely <20°C), the disease was less severe. 3. in a third period with identical conditions (high RH and high temperatures) the disease again became severe. The observations indicate that there are interactions between weather variables which need to be taken into account to explain disease dynamics.

Another potential driver of disease dynamics is the change in leaf susceptibility with time. The rubber tree grows by a succession of flushes, either all year round, when the tree is immature, or during a process of defoliation-refoliation, when the old leaves fall and are replaced by new leaves at a precise period of the year. Only immature leaves are receptive to SALB and their receptivity decreases as they mature. The immature period of each leaflet development lasts 15–21 days (Guyot et al. 2008), while the duration of flush development, from budburst to complete maturity, is about one month. Rubber tree growth is rhythmic in about the first eight years, due to the succession of dormancy and growth phases, depending on genetic and environmental factors. In the first six months after planting, or after complete pruning, the young rubber tree can produce one or two flushes quarterly, which are somewhat synchronised between trees. Once the tree begins to branch laterally, this synchronicity disappears and each tree can produce new flushes uninterruptedly as long as rainfall is sufficient. During the dry season, flush production tends to decrease and to stop until the rainy season returns. Older trees become mature, with a cyclic phenology characterised by an annual period of defoliation and refoliation. Given the particular phenological behaviour of the rubber tree, its consequences for leaf diseases have to be considered on five scales: the leaf, the flush, the tree, the monoclonal plot and the polyclonal plantation.

In this paper, the influence of the climate on SALB severity was studied for both flush and leaf scales: the climatic variables exerted over each flush were related to disease severity on that flush and the climatic variables exerted over each development stage of each leaflet were related to disease severity on the same leaflet.

The aim of this study was to identify the weather variables that have, individually or in interaction, the greatest influence on SALB dynamics under Amazonian conditions, and to predict the damage risks based on knowledge of the climatic variables, taking into account changes in host leaf susceptibility. The analysis was performed with classical methods and with the CART method (Classification And Regression Trees) developed by Breiman et al. (1984). Although it is unusual in plant epidemiology, CART is frequently used in human health studies (Dorel and Perrier 1990; Hu et al. 2008; Duc Tang et al. 2008; Muller and Möckel 2008), the social sciences and for risk assessment, especially in industry (Chang and Chen 2005; Bevilacqua et al. 2008). De'Ath and Fabricius (2000) also used CART for an ecological

data analysis in a coral reef. This non-parametric procedure does not need *a priori* hypotheses on the data set and can reveal interactions between the explicative variables, does away with linear or additive relations, and can cope with missing data. Furthermore, presentation of the results in tree form makes for easy reading and interpretation of the analysis. This segmentation method allows both discriminant analyses and multiple regressions using both quantitative and qualitative variables as the predictive variables.

The CART method is, therefore, a powerful alternative to conventional discriminant methods, such as linear discriminant analysis and logistic regression, and to regression methods, such as multiple linear regressions. Compared with the latter method, CART provides results (regression trees) directly derived from the original data; using non-parametric and non-linear techniques, the method is more suitable for addressing complex problems. Finally, large arrays of explanatory variables, even when correlated to each other, can be handled and their interactions are accounted for, provided there are enough data to obtain stable decisions trees with a significant number of individuals per node (Gray and Fan 2008). For the present analysis, a wide array of weather variables was considered to highlight the most relevant ones for disease development. CART therefore remains an exploratory method and, in the literature, it is always used as a complement to other methods.

CART creates a binary tree to analyse a variable (here, the disease index) with a set of explanatory variables (here, the weather variables) (Nakache et al. 1996). For continuous variables, such as the disease index, regressions are used to build the tree.

From a set of predictors, CART splits the original data set (called the root node) into two subsets (called child nodes) so that the values of the explained variable in each child node are as similar as possible and as different as possible between the two child nodes. For the next partition, each child node becomes a parent node and is divided again into two child nodes. From one step to the others, the homogeneity in the nodes increases. Once a partition is constructed, it is not reconsidered for the subsequent partitions. Thus, each branch of the tree is processed independently, so that the interactions between predictors are more effectively taken into account than in a conventional regression process. To accomplish this partitioning, CART evaluates all the possible splits with all predictors and chooses the best split which corresponds, of course, to the best explicative predictor. The best split is determined with 'impurity' criteria whose calculations are described by Berk (2006). CART discards any statistical rule to determine the end of segmentation, so that division can go on as soon as possible. Consequently, it is necessary to define a criterion to stop the divisions, for example the minimum number of data in a node or the number of steps of the tree. The succession of dichotomies leads to the formation of a segmentation tree which needs pruning in order to collapse terminal nodes that do not reduce heterogeneity sufficiently for the extra complexity added (Berk 2006). Pruning is a trade-off between classification errors and the complexity of the tree.

Materials and methods

Experimental set-up

The trial was set up as two small designs, planted at high densities in Pointe Combi (5°20′N–52°55′W), French Guiana. Annual rainfall in the area is 3,000 mm with a dry season from August to November (Fig. 1). The climatic conditions registered during the survey of design A (Figs. 2 and 3) were representative of the overall climatic pattern in French Guiana: high RH throughout the year and very little fluctuation in temperatures.

Design A, intended to analyse disease dynamics on a flush scale, comprised a single 100 m^2 plot, with 100 trees of the susceptible clone IRCA GY 5 planted 1 m×1 m apart. Observations were made from December 2005 to June 2007 on 412 flushes.

Design B was intended to compare the effect of climate on host-parasite interactions on clones with different susceptibilities to the disease, and to take into consideration the incidence of leaf development on the climate effect. It included three clones differing in their level of resistance to SALB (Guyot et al. 2008): IRCA GY 5 (susceptible), PB 260 (susceptible) and FX 3864 (partially resistant). For each clone, 60 trees were planted in a 7.5 m×4.5 m plot, with four rows, each containing 15 trees. The plots were 10 m apart. To keep the trees short enough for observations,

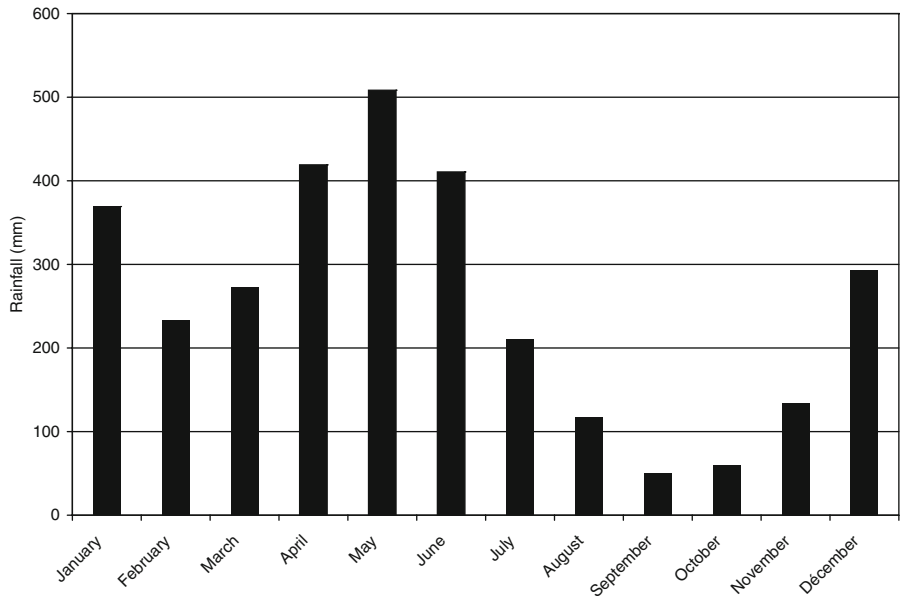

Fig. 1 Average monthly rainfall (mm) recorded at Pointe Combi (French Guiana), 1980–2007

and to have permanently young leaves, five trees per clonal plot were cut back each month. Observations were made from November 2000 to December 2002 on 2,990, 3,846 and 4,222 leaves for clones IRCA GY 5, PB 260 and FX 3864, respectively.

Phenology assessment

One flush per tree in design A and 1–5 flushes per clone in design B were simultaneously observed, from budburst to the maturity of the youngest leaf. In design A, it was recorded each week whether the flush was immature or mature. In design B, the development status of each leaflet was recorded every three days.

The leaf stages were identified according to the scale of Hallé and Martin (1968), modified as follows: B1 (duration 3–6 days, leaflets folded dorsally, pointing upward, reddish), B2a (duration 4–6 days, leaflets partially or totally unfolded, pointing downward, shiny reddish), B2bC (duration 5–8 days, leaflets unfolded, pointing downward, dull and light reddish, then light green), and D (maturity, leaflets unfolded—horizontal, shiny green with hard lamina).

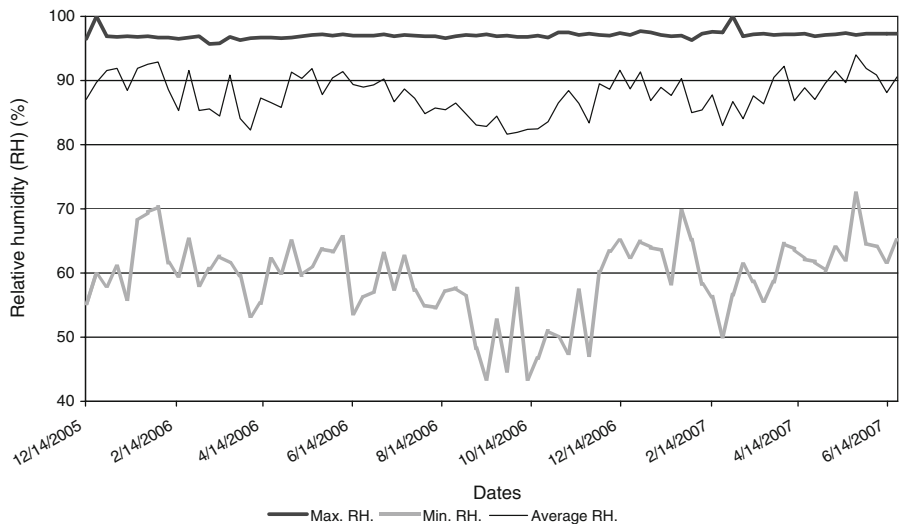

Fig. 2 Weekly maximum, minimum and average relative humidity (RH), as a %, recorded in design A, December 2005–June 2007

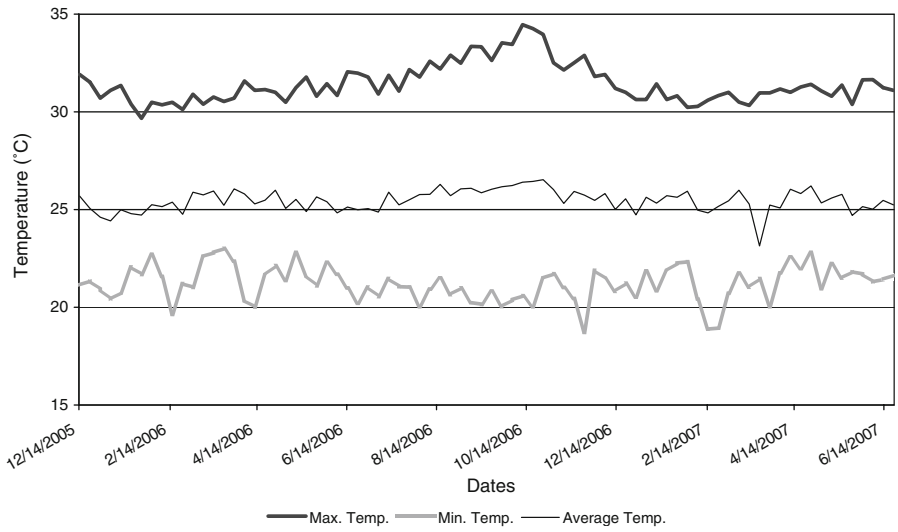

Fig. 3 Weekly temperatures (maximum, minimum and average) in °C recorded in design A, December 2005–June 2007

Disease assessment

When all the leaves in a flush were mature, disease severity was evaluated for each leaflet, with a disease index (DI) that characterised the necrotised area of the lamina according to Chee's scale (1976a) modified as follows: 0 = nil; 1 = low (<1% leaf area necrotised); 2 = medium (1–10% leaf area necrotised); 3 = high (10–30% leaf area necrotised); 4 = very high (large necrosis, >30% leaf area necrotised); 5 = leaflet fall due to SALB (Guyot et al. 2008). On a flush scale, the disease severity index (DSI) was the average of the disease indices (DI) of its leaflets. Both DI and DSI were scored as low (L), medium (M) or high (H) (Table 1).

Measurement of weather variables

In design A, the following weather variables were recorded every 2 min with an automatic Campbell CR10X station with the following sensors: HMPC 45 for the temperature and the RH, 237F for wetness, and a tipping bucket rain gauge (tipping every 0.2 mm) ARG100 (Campbell Scientific Ltd., Shepshed, United Kingdom). The incoming global radiation (sun plus sky, waveband 300–2500 nm) was measured with a Cimel Enerco 47 station (Cimel Électronique, Paris, France) and a pyranometer CE 180. In design B, the climate was measured continuously. The rainfall was recorded with a pluviograph equipped with a tipping bucket, while temperature and RH was measured with a thermohygrograph, also equipped with a wetness sensor. In both designs, the wetness sensors were placed vertically so that their positions were identical to that of the laminae of immature leaflets. The whole immature periods of each flush in design A and each development stage of each leaflet in design B were characterised by the weather conditions encountered during the period (Table 1).

Statistical analysis

The statistical analysis was completed using classical methods with SAS 9.1. software for the leaflet models (Pearson's linear correlations and multiple linear regressions) and the CART method for both flush and leaflet models. The CART method was used to explore the relationships between the disease index and the weather variables, without any limitation to linear relationships, and to bring out possible interactions between weather variables. It was decided to stop the segmentation when one of the following conditions was reached: four levels under the root node or minimum improvement of impurity of 0.0001, or minimum number of data in a node of five on the flush scale (design A) and 50 on the leaflet scale (design B). The software AnswerTree (SPSS Inc., Chicago, Illinois, USA), for ordinal data on the leaflet scale, was used. The Gini index (http://www.answers.com/topic/decision-tree-learning) was used to estimate impurity, and pruning was based on the standard error rule.

Table 1 Summary of the climatic variables and disease severity index classes used in the flush and leaflet models

	Flush model	Leaflet model
Incoming global radiation	+	
Average temperature	+	+
Average temperature between noon and 4p.m. (warmest period of the day)	+	+
Average temperature between midnight and 4a.m. (coolest period of the day)	+	+
Average relative humidity (RH)	+	
Average diurnal relative humidity (RH)		+
Average nocturnal relative humidity (RH)		+
Average relative humidity (RH) between noon and 4p.m. (driest period of the day)	+	
Average relative humidity (RH) between midnight and 4a.m. (most humid period of the day)	+	
Duration of relative humidity was ≥90%	+	
Duration of relative humidity was ≥95%	+	
Total rainfall	+	+
Total rainfall duration	+	+
Total wetness duration	+	+
Maximal uninterrupted wetness duration		+
Disease Severity Index: low (L)	0≤DI≤1	DI=0 or DI=1
Disease Severity Index: medium (M)	1<DI≤3	DI=2 or DI=3
Disease Severity Index: high (H)	3<DI≤5	DI=4 or DI=5

Results

Correlations and multiple regressions on a leaflet scale

Although the Pearson's correlation coefficients were highly significant because of the very large number of data, they were considered as weak (Table 2). The highest Pearson's correlation coefficient was 0.38 (between maximum uninterrupted wetness duration and disease severity on IRCA GY 5). The coefficients of determination of the multiple linear regressions were also low (Table 3) and the multiple linear regressions could, in the best cases, explain only 16–40% of the variations in disease severity with the climatic variables. Several models were tested for each clone in order to avoid too great a number of variables and to discard variables with too high a risk of interactions. The models were found to be very unstable depending on the variables introduced. The different transformations tested (inverse, square root, logarithmic, inverse of square root), likewise the graphs, did not improve the results and did not show a clear relationship between disease severity and the climatic variables under the conditions of our study.

CART analysis on a flush scale

On a flush scale (Fig. 4), on the susceptible clone IRCA GY 5, the effects of climatic conditions were recorded over the entire duration of flush growth. The first segregating variable was the average RH. When it exceeded 89% (node 2), most of the flushes were highly diseased, but many flushes with DSI = 3 were located in node 1, which was characterised by an average RH <89%. The second segregating variable, the average temperature, split node 1 into two child nodes with a threshold at 25°C. Most of the highly diseased flushes were in node 4 (average temperature >25°C). Nodes 7 and 8 indicated that very abundant rainfall was detrimental to SALB. Nodes 9 and 10 showed that the DSI was somewhat medium when the RH during the driest hours of the day was <60%.

CART analysis on a leaflet scale

For IRCA GY 5 (Fig. 5), the most discriminatory variable was the total wetness duration in B2a. When it was >9 h during the total duration of stage B2a, the DI was high (node 2). Such a situation was very common and concerned 75% of the leaflets. During

Table 2 Pearson's correlations between climatic variables and disease index on flush and leaflet scales

Climatic variable	Leaflet stage	Coef. correl.	Proba.	Nb of data
Flush				
Average relative humidity (RH)		0.33	<0.0001	412
Average relative humidity (RH) between noon and 4p.m.		0.27	<0.0001	412
Average relative humidity (RH) between midnight and 4a.m.		0.23	<0.0001	412
IRCA GY 5 leaflets				
Maximal uninterrupted wetness duration	B2a	0.38	<0.0001	1991
Total wetness duration	B2a	0.37	<0.0001	1991
Average diurnal relative humidity (RH)	B1	0.26	<0.0001	2555
PB 260 leaflets				
Total wetness duration	B2a	0.35	<0.0001	2527
Average relative humidity (RH)	B2bC	0.34	<0.0001	3547
Average diurnal relative humidity (RH)	B2bC	0.31	<0.0001	3547
FX 3864 leaflets				
Maximal uninterrupted wetness duration	B2bC	−0.23	<0.0001	2862
Maximal temperature	B1	0.21	<0.0001	3738
Maximal temperature	B2bC	0.21	<0.0001	4160

the survey of design B, the wetness duration was >9 h for 92% of the days. About two thirds of the leaflets subjected to >9 h of wetness duration in B2a were severely necrotised. This represented 86% of the leaflets with DI=3. However, 2.4% of the leaflets in node 2 had a low DI and the leaflets with DI=1 were distributed approximately equally between nodes 1 and 2. This indicated that the wetness duration was not very discriminatory for light damage. Rainfall duration and temperature were respectively the second and the third discriminatory variables. Very long rainfall in B1 appeared to be detrimental to SALB while moderately long rainfall in B2bC favoured low and high DI. Average temperatures <24–25°C were not favourable to SALB (nodes 7 and 9) but a nocturnal temperature <24°C in B1 encouraged the disease (node 11). The terminal nodes highlighted the most suitable and the least suitable conditions for SALB. Node 9 (total wetness duration in B2a≤9 h rainfall duration in B2bC>6 h average temperature in B1>24°C) represented the conditions leading to light damage. The climatic conditions of node 11 (total wetness duration in B2a>9 h rainfall duration in B1≤ 37 h average nocturnal temperature in B1≤24°C) were the most favourable to SALB. These observations proved that unfavourable conditions for SALB could be balanced by favourable conditions and lead to severe disease. The inverse situation could also occur. The segmentation tree also indicated that the conditions prevailing at all development stages influenced the disease.

For PB 260 (Fig. 6), as for IRCA GY 5, the most important variable was the wetness duration in B2a, here, the uninterrupted wetness duration. The threshold was higher (13.5 h) than for IRCA GY 5 and the leaflets subjected to the longer uninterrupted wetness duration had a high probability (0.86) of being severely diseased (node 2). However, such conditions were relatively infrequent and they concerned only 29.5% of the leaflets. Most of the leaflets with DI=3 were located in node 1, i.e. with an uninterrupted wetness duration under 13.5 h. The average nocturnal RH in B2bC split node 1 into two nodes with a threshold of 94%. A high average nocturnal RH was very conducive to SALB. When, in addition, the average RH was >85%, the DI was high for 67% of the leaflets. Thus, node 8 was considered as representative of the most suitable conditions for SALB. Conversely, the leaflets with low DI were distributed in several nodes characterised by different climatic conditions. Node 10 was the node where the leaflets had the greatest probability of avoiding severe disease (0.88) but node 6 was the node which contained the largest number of leaflets with DI=1.

For the resistant clone FX 3864 (Fig. 7), <5% of the leaflets showed a high DI. The uninterrupted

Table 3 Multiple linear regressions between climatic variables and disease index for leaflet models

Predictor variable	Parameter estimate	$P>\|t\|$	Predictor R^2	Model R^2	Model *F value*	Model $P>\|t\|$
Flush disease index model						
Intercept	−53.96	<0.0001				
Average RH	0.28	<0.0001	0.11			
Average temperature	1.26	<0.0001	0.05	0.19	23.66	<0.0001
Incoming global radiation	1.50E−05	0.0013	0.02			
Total rainfall	1.11E−04	0.0372	0.01			
IRCA GY 5 leaflet Disease Index model						
Intercept	−20.41	<0.0001				
Average noon—4p.m. temperature in B2a	−0.51	<0.0001	0.16			
Uninterrupted wetness in B2bC	−0.08	<0.0001	0.06	0.26	20.97	<0.0001
Rainfall duration in B2bC	0.00122	<0.0001	0.02			
Average diurnal RH in B2bC	−0.33	<0.0001	0.02			
PB 260 leaflet Disease Index model						
Intercept	−22.91	<0.0001				
Average RH in B2bC	5.59	<0.0001	0.23			
Average nocturnal RH in B2a	0.33	<0.0001	0.11	0.40	74.82	<0.0001
Average temperature in B2a	0.14	<0.0001	0.04			
Average noon—4p.m. temperature in B1	−1.59	<0.0001	0.03			
FX 3864 leaflet disease index model						
Intercept	17.35	<0.0001				
Uninterrupted wetness in B2a	−0.03	<0.0001	0.06			
Average RH in B1	0.97	<0.0001	0.04	0.16	25.70	<0.0001
Average temperature in B2bC	2.19	<0.0001	0.04			
Average diurnal RH in B1	−0.46	<0.0001	0.02			

wetness duration was also the predominant variable in the segregation, but on stage B2bC. However, although 73% of the leaflets subjected to an uninterrupted wetness duration of >13 h exhibited a DI=1 (node 2), the leaflets with that low DI were approximately equally distributed between nodes 1 and 2. The major difference concerned the medium DI, which was most frequent in node 1, as a percentage of the leaflets in the node and among all leaflets with DI=2. The subsequent variables in the segregation were diversified (rainfall, average, diurnal or nocturnal temperature, diurnal RH, wetness duration) and applied to all the leaflet development stages, so that it was difficult to place the influence of the climatic variables in order of importance for this clone. However, node 9 (uninterrupted wetness duration in B2bC>13 h, average diurnal RH in B1> 89% and total wetness duration in B2a≤24 h) revealed the most propitious climatic conditions for SALB. Conversely, the conditions of node 5 (uninterrupted wetness duration in B2bC>13 h, average diurnal RH in B1≤89%) appeared to be the most detrimental conditions for SALB (DSI=1 represented 78.5% of the leaflets of the node and 47.5% of all the leaflets with DI = 1). On this clone, the presence of free water did not seem to be an essential condition for severe disease.

Discussion

Few authors have used the CART method in plant epidemiology. Baker et al. (1993) evaluated the hazard of pine mortality caused by *Heterobasidion annosum* with 22 soil measurements as candidate predictors. De Lapeyre de Bellaire and Dubois (1997) assessed the distribution of thiabendazole-resistant *Colletotrichum musae* isolates from banana planta-

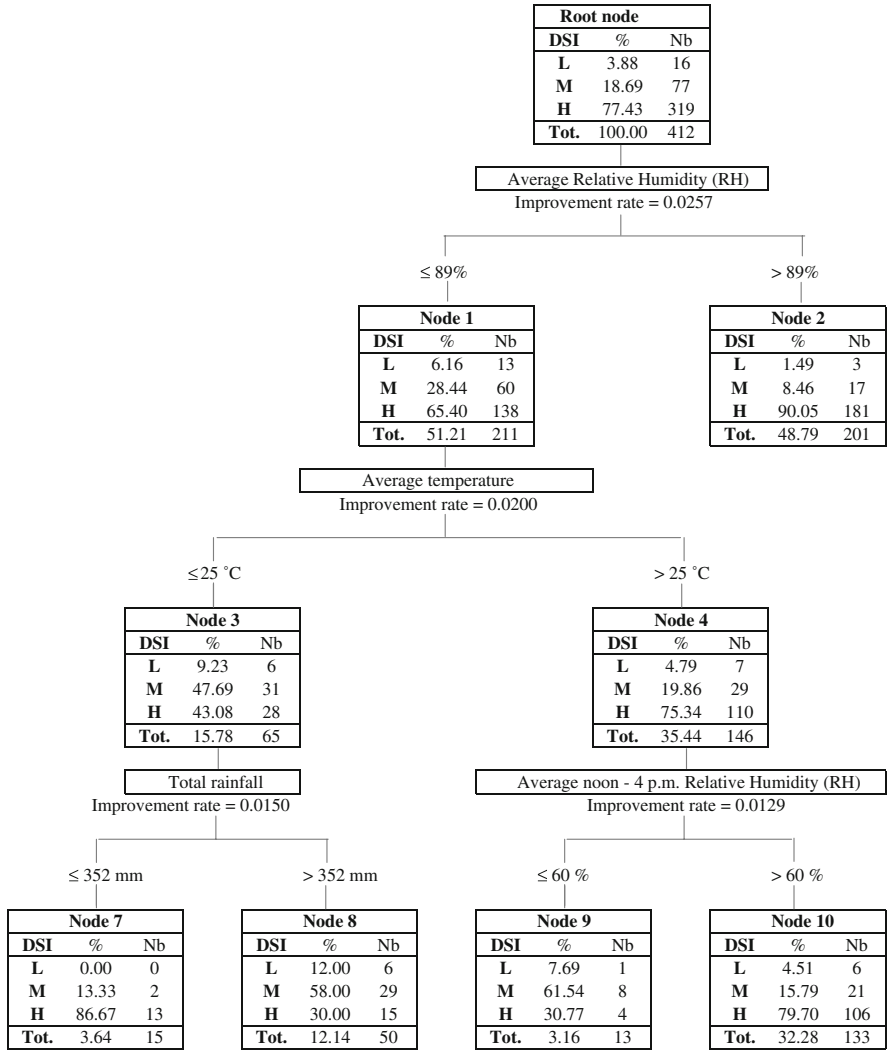

Fig. 4 Segmentation tree for the effect of weather variables on the disease index on a flush scale for susceptible clone IRCA GY 5. For each segmentation, segregating variable, improvement rate, segregating value of the variable; for each node, percentage and number of each DSI: low (L), medium (M), high (H), number of leaflets in the node and percentage of these numbers compared with the number of leaflets observed for clone IRCA GY 5

tions according to treatment procedures and climate variables; (De Lapeyre de Bellaire 1999) used CART, among other classical methods, to explore the effects of the diversity of soil characteristics and cultural practices on the level of banana fruit contamination by *C. musae*. CART was also used to detect and classify limiting factors for banana cropping (Perrier and Delvaux 1991) and to determine the influence of the environment and of cultural practices on banana yield (Dorel and Perrier 1990). Avelino (1999) and Avelino et al. (2006) proposed a CART segmentation analysis to characterise the factors related to coffee rust in Honduras. In the present study, use of the CART method shed light on the relations existing between the climate and the disease, which classical methods were unable to reveal. It made it possible to place climatic variables in order of importance according to their influence over the disease and to define, for the main variables, limits beyond which their action was rather positive or rather negative. However, CART was unable to assess the explanatory potential of the model. It will be seen later that the conclusions provided by CART were perfectly compatible with the data in the literature.

CART showed cause and effect relations between the climate and SALB, which conventional linear regressions or correlation calculations barely bring out. Of course, this analysis method was particularly adapted to this study because the limited climatic variations meant that it was not possible to bring out a variable limiting the disease. During the present study, this technique was found to enable quite easy

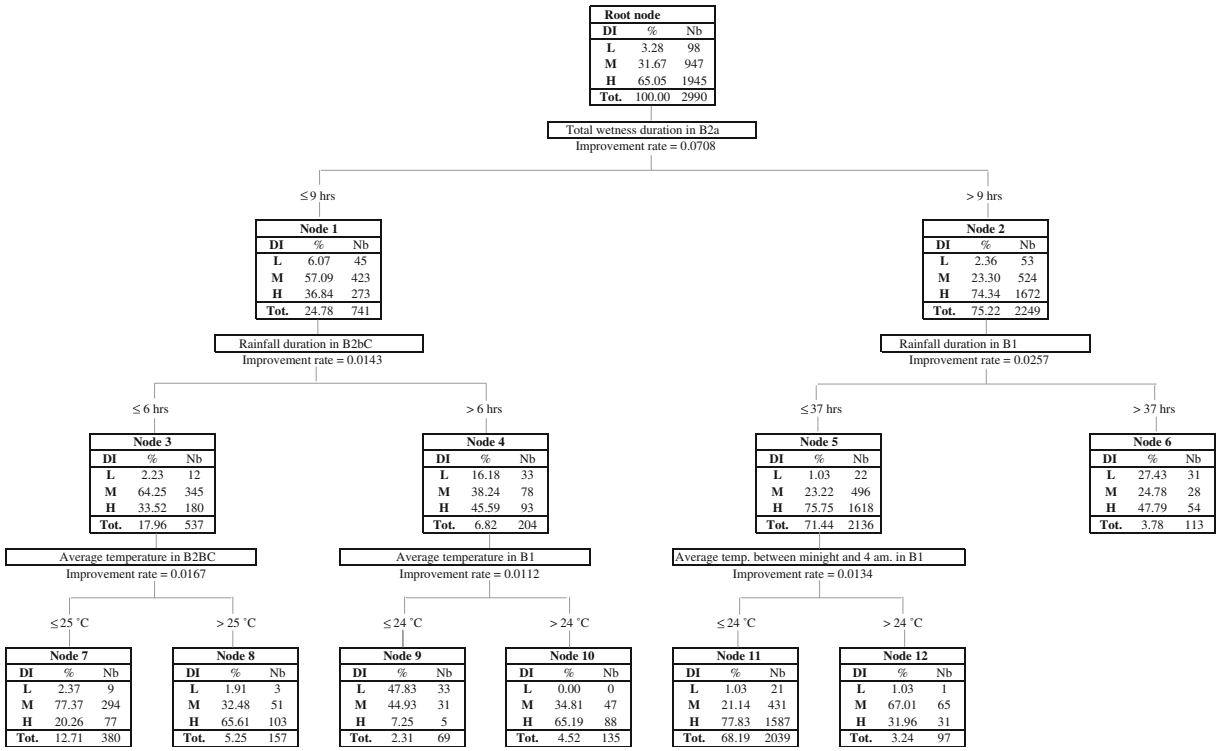

Fig. 5 Segmentation tree for the effect of weather variables on the Disease Index (DI) on a leaflet scale for the susceptible clone IRCA GY 5. For each segmentation, segregating variable, improvement rate, segregating value of the variable; for each node, percentage and number of each DI: low (L), medium (M), high (H), number of leaflets in the node and percentage of these numbers compared with the number of leaflets observed for clone IRCA GY 5

prediction of the variable to be analysed according to the predictive values, but it was much less suitable for explaining the disease as a function of climate. Gray and Fan (2008) indicated that one of the drawbacks of CART is the instability of its solutions. However, such instability was found in multiple linear regressions. In order to reduce that problem, it was chosen to work on a large number of data, so there was no need to validate the tree by another series of data.

It still remains that, as for the other methods moreover, the solutions proposed by CART were only valid under the study conditions or similar conditions. Such an analysis carried out under very different climatic conditions would probably give equally very different results, not because of the method but because of different interactions between variables and the effect of the climate on the disease. In our study, as in the work by Baker et al. (1993), the classical methods of analysis did not bring out strong linear relationships between the climate and the disease, and the influence of small variations in the weather variables on disease development could not be demonstrated by simple mathematical relations in such a stable environment. The CART method highlighted the weather variables that most influence infections by *M. ulei* in the Amazonian environment. Moreover, the CART method considered the interactions between weather variables as drivers of the disease and detected non-linear relationships. The CART method enabled us to define the combinations of weather variables conducive to different levels of disease, taking into account the clonal susceptibility of the host.

The two trials showed that, once the disease was present in a plot and when the trees produced young receptive leaves, the climatic conditions were never limiting for infections and the development of necrosis, be it for susceptible clones or resistant clones, under our experimental conditions. Despite these differences between clones, some similarities clearly emerged in the results. It was obvious that high RH for long periods was the most important

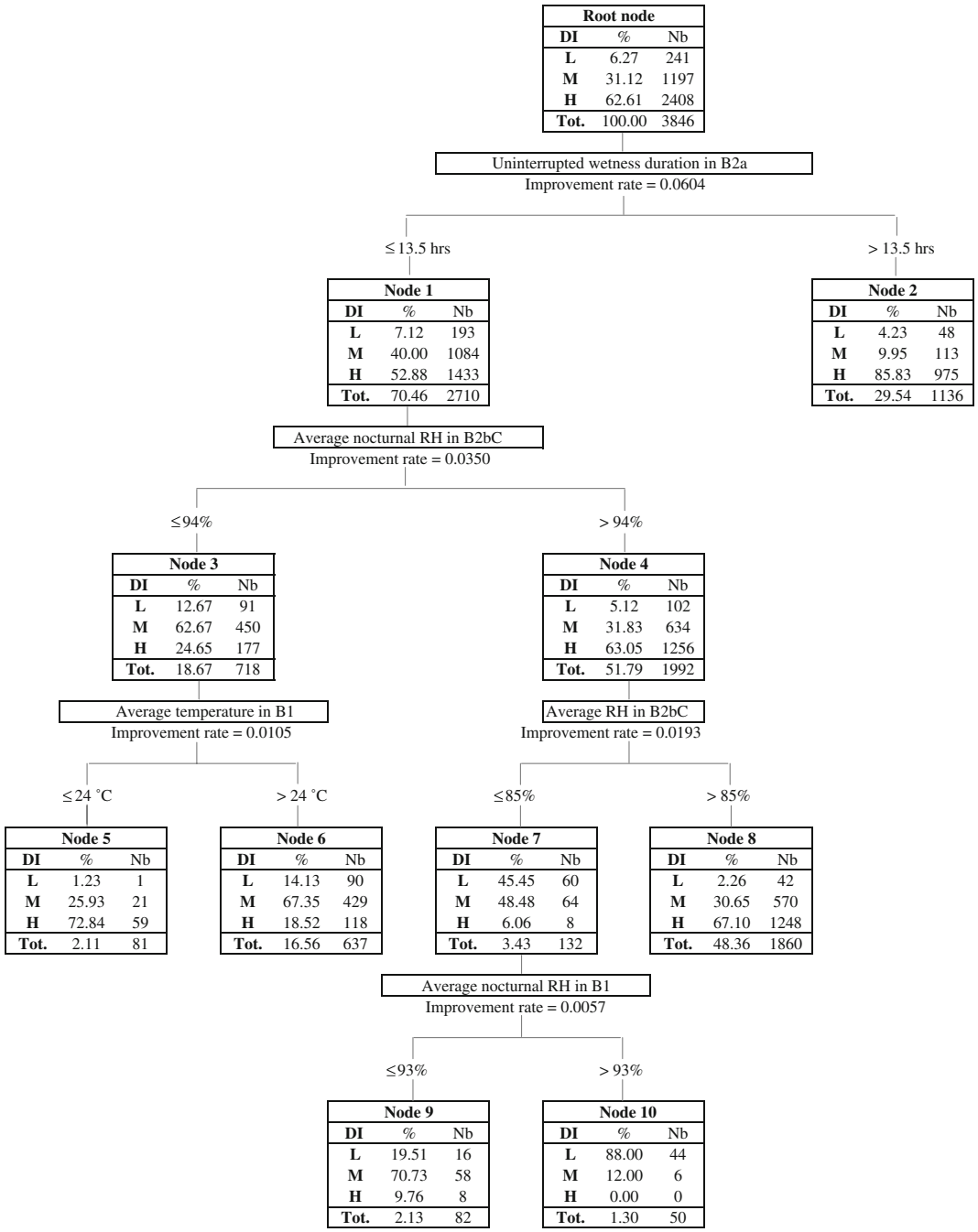

Fig. 6 Segmentation tree for the effect of weather variables on the Disease Index (DI) on a leaflet scale for the susceptible clone PB 260. For each segmentation, segregating variable, improvement rate, segregating value of the variable; for each node, percentage and number of each DI: low (L), medium (M), high (H), number of leaflets in the node and percentage of these numbers compared with the number of leaflets observed for clone PB 260

factor for severe disease. Such conditions occur very often in Amazonia: during our trials, a period of >9 h during which RH was >95% was encountered for 59% of the days. Long wetness duration and average temperatures over 25°C were also conditions highly conducive to the disease in the field. Such climatic conditions were very frequent: the wetness duration was >9 h for 92% of the days. Under controlled

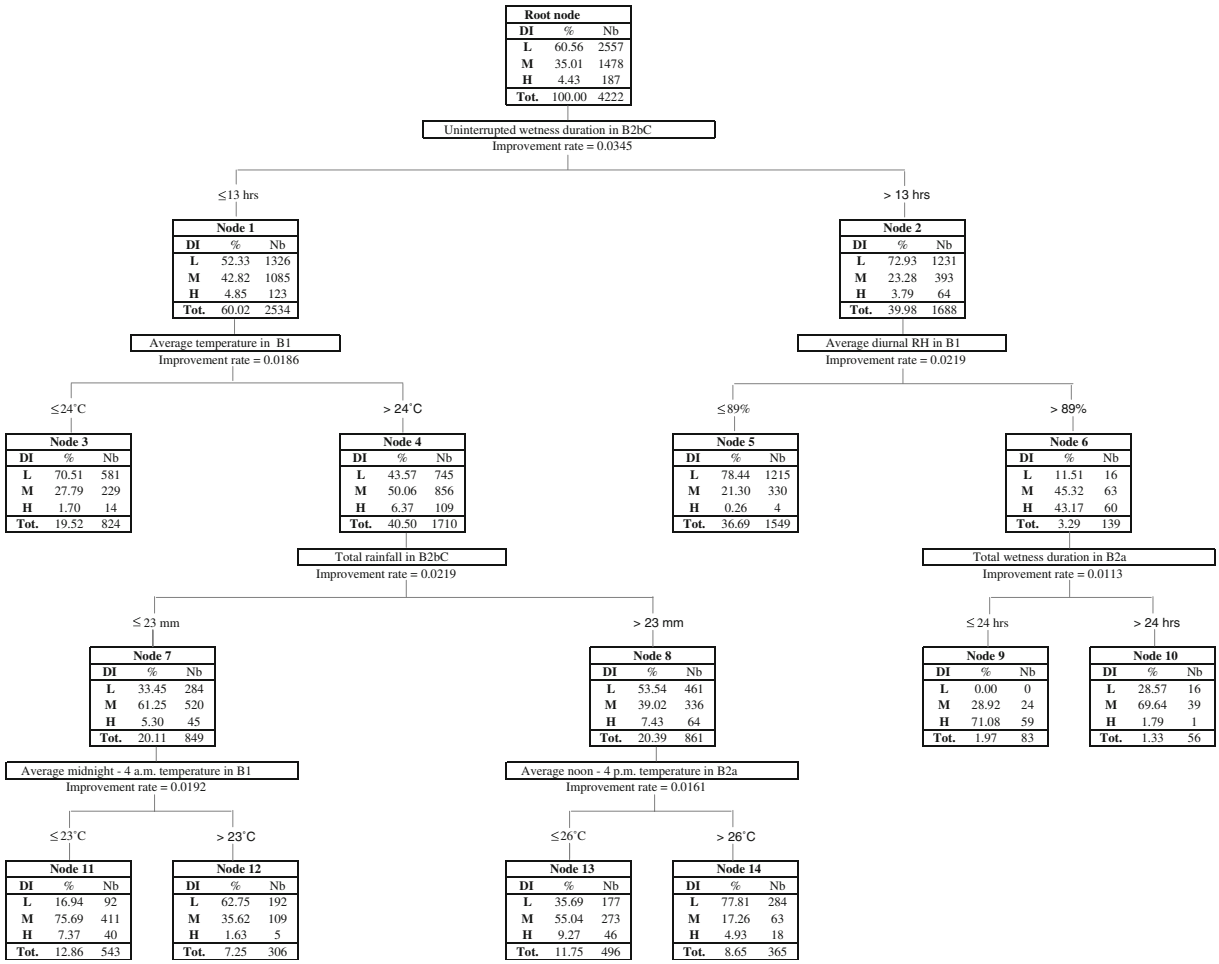

Fig. 7 Segmentation tree for the effect of weather variables on the Disease Index (DI) on a leaflet scale for the partially resistant clone FX 3864. For each segmentation, segregating variable, improvement rate, segregating value of the variable; for each node, percentage and number of each DI: low (L), medium (M), high (H), number of leaflets in the node and percentage of these numbers compared with the number of leaflets observed for clone FX 3864

conditions, the optimum for *M. ulei* germination and growth was between 24°C and 28°C. Under the very high RH prevailing in the Amazonian region, the amount or duration of rainfall were not essential predictors for the disease and excess water even seemed unfavourable to SALB: the percentage of low diseased leaflets or flushes was higher for >23 mm in B2bC for FX 3864, >6 h of rainfall in B2bC for IRCA GY5 and >352 mm during the whole flush development in IRCA GY 5. All these results agreed with those obtained by Gasparotto et al. (1989b) at Ponte Nova. The climatic conditions were influential at every development stage.

For the three clones, the climatic conditions prevailing at all development stages were important for the final disease index. This suggested that the disease index resulted from two biological processes, infection, requiring inoculum availability, and successful fungus penetration and necrotic expansion. Thus, unsuitable conditions at one stage of leaflet development could be compensated for by suitable conditions at another stage, and vice-versa. The comparison between clones on a leaflet scale showed that it was not the same weather variables that were the most discriminant for the three clones. For IRCA GY 5, wetness, then rainfall and lastly temperatures most explained the variations in disease index. For the other susceptible clone, PB 260, rainfall did not play a major role. For the resistant clone, FX 3864, wetness duration, RH, rainfall and temperature greatly influ-

enced disease severity. In all cases, wetness duration was the most important variable for disease severity. It also seemed that the same variable could act in opposite ways depending on the leaf stage. For example, this was the case for continuous wetness on clone FX 3864. A similar observation was reported by Mishra and Bhattacharyya (2001) who found that, depending on the developmental stages of banana, the climate had a positive or negative effect on *Mycosphaerella musicola* attacks.

The fact that the conditions subjected to in B1 had an effect on the disease index suggests that infections could occur at that stage, for which susceptibility has never been demonstrated or even assessed. Most pathogenic fungi are not very sensitive to outside RH conditions after penetration (Agrios 2005). Yet, the positive effect of rainfall in B2bC on IRCA GY 5, the effect of RH at that stage on the three clones and that of wetness in B2bC on resistant clone FX 3864 lead to two hypotheses: either *M. ulei* is a fungus that is sensitive to outside conditions once it has penetrated its host, or infections can occur at stage B2bC, under field conditions, in sufficient numbers to affect the disease index. Yet stage B2bC is the stage from which the host acquires its resistance and is considered not very susceptible to infections. The occurrence of ontogenic resistance therefore needs to be studied, under both controlled and natural conditions. It should involve several clones with differing degrees of susceptibility. Indeed, Junqueira et al. (1990) found that the periods of rubber tree leaf receptivity to *M. ulei* varied from 10 to 16 days depending on the clone and that those differences had consequences for sporulation and development of sexual structures. These results suggest that ontogenic resistance can occur at different times in different clones.

The results of this study show the difficulty in precisely defining conditions conducive to the disease when there is no limiting environmental factor, due to the fact that numerous interactions exist between weather variables and the final disease index results from the conditions subjected to at all immature leaf stages. This clearly highlights the risks of error that can occur from the extrapolation of results obtained under controlled conditions. Field experiments are essential for predictive SALB models based on climate. Indeed, it may affect the phenology of the host, the ability of the parasite to sporulate, or the success of infections and necrosis expansion. It can also affect synchronisation between host receptivity and inoculum availability. It is therefore difficult to define escape zones in a predictive manner and only observations of all the components of the disease in locations displaying varied climatic conditions can help to specify escape conditions and explain how they act upon disease.

Lastly, this study only focused on infection and necrosis development, given that the disease was present in the plot and the trees displayed active vegetation. It did not take into account the conditions needed to trigger the epidemic and for the existence of young receptive leaves.

Acknowledgements The authors would like to thank Mr. Peter Biggins for the English translation of this article.

References

Agrios, G. N. (2005). Brown rot of stone fruits. *In: Plant pathology, 5th edition, Elsevier*, p. 254.

Avelino, J. (1999). Contribution à la gestion du pathosystème *Hemileia vastatrix* et *Coffea arabica* au Honduras. Influence des principaux facteurs physiques, phytotechniques et biologiques. *Thèse de Doctorat. Université de Paris XI, Orsay, France*, 127 pp.

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. *Ecological Modelling, 197*, 431–447.

Baker, F. A., Verbyla, D. L., Hodges, C. S., & Ross, E. W. (1993). Classification and regression tree analysis for assessing hazard of pine mortality caused by *Heterobasidion annosum*. *Plant Disease, 77*(2), 136–139.

Berk, R. A. (2006). An introduction to ensemble methods for data analysis. *Sociological Methods & Research, 34*(3), 263–295.

Bevilacqua, M., Ciarapica, F. E., & Giacchetta, G. (2008). Industrial and occupational ergonomics in the petrochemical process industry: a regression trees approach. *Accident Analysis and Prevention, 40*, 1468–1479.

Breiman, L., Friedman, J. H., Olshen, R., & Stone, C. J. (1984). *Classification and regression trees*. Belmont: Wadsworth.

Chang, L. Y., & Chen, W. C. (2005). Data mining of tree based-models to analyse freeway accident frequency. *Journal of Safety Research, 36*, 365–375.

Chee, K. H. (1976a). Assessing susceptibility of *Hevea* clones to *Microcyclus ulei*. *Annals of Applied Biology, 84*, 135–145.

Chee, K. H. (1976b). Factors affecting discharge, germination and viability of spores of *Microcyclus ulei*. *Transactions of the British Mycological. Society, 66*(3), 499–504.

Chee, K. H. (1980). The suitability of environmental conditions in Asia for the spread of South American leaf blight of *Hevea* rubber. *Planter, 56*, 445–454.

Chee, K. H. (1985). An analysis of possible preventive measures against the introduction of South American leaf blight to Malaysia. *Plant Quarantine Support for Agricultural Development Regional Conference, Kuala Lumpur. December 1985*, 261–263

Chee, K. H., Darmono, T. W., Zhang, K. M., & Lieberei, R. (1985). Leaf development and spore production and germination after infection of *Hevea* leaves by *Microcyclus ulei*. *Journal of the Rubber Research Institute of Malaysia, 33*(3), 124–127.

De'Ath, G., & Fabricus, K. E. (2000). Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology, 81*(11), 3178–3192.

De Camargo, A. P., Cardoso, R. M. G., & Schmidt, N. C. (1967). Comportamento e ecologia do "mal-das-folhas" da seringueira nas condiçoes climaticas do planalto Paulista. *Bragantia, 26*(1), 1–8.

De Lapeyre de Bellaire, L. (1999). Bio-écologie de *Colletotrichum musae* (Berk. & Curt.) Arx, agent de l'anthracnose des bananes, dans les conditions tropicales humides de la Guadeloupe. *Thèse de Doctorat, Université de Paris XI, Orsay, France*, 100 pp.

De Lapeyre de Bellaire, L., & Dubois, C. (1997). Distribution of Thiabendazole-Resistant *Colletotrichum musae* isolates from Guadeloupe Banana Plantations. *Plant Disease, 81* (12), 1378–1383.

Dorel, M., & Perrier, X. (1990). Influence du milieu et des techniques culturales sur la productivité des bananeraies de Guadeloupe. Enquête-Diagnostic. *Fruits, 45*(3), 237–244.

Duc Tang, N., Erhart, A., Speybroeck, N., Xuan Hung, L., Khanh Than, L., Trinh Hung, C., et al. (2008). Malaria in central Vietnam: analysis of risk factors by multivariate analysis and classification tree models. *Malaria Journal, 7* (28), 1–9.

Garcia, D., Mattos, C. R. R., de Gonçalves, P. S., & Le Guen, V. (2004). Selection of rubber clones for resistance to South American leaf blight and latex yield in the germplasm of the Michelin Plantation of Bahia (Brazil). *Journal of Rubber Research, 7*, 188–198.

Gasparotto, L. (1988). Epidemiologia do mal das folhas (*Microcyclus ulei* (P. Hem) v. Arx) da seringueira (*Hevea* spp.). *PhD thesis, Universidade Federal, Viçosa (Minas Gerais, Brasil)*, 124 pp.

Gasparotto, L., Zambolim, L., Maffia, L. A., Ribeiro do Vale, F. X., & Junqueira, N. T. V. (1989a). Efeito da temperatura e umidade sobre a infecçaõ de Seringueira (*Hevea* spp.) por *Microcyclus ulei*. *Fitopatologia Brasileira, 14*(1), 38–41.

Gasparotto, L., Zambolim, L., Ribeiro do Vale, F. X., Maffia, L. A., & Junqueira, N. T. V. (1989). Epidemiologia do mal das folhas da seringueira. I. Ponte nova-MG. *Fitopatologia Brasileira, 14*(1), 65–70.

Gasparotto, L., Zambolim, L., Junqueira, N. T. V., Maffia, L. A., & FX, Ribeiro do Vale. (1991). Epidemiologia do mal das folhas da seringueira. II. Regiaõ de Manaus-AM. *Fitopatologia Brasileira, 16*(1), 18–21.

Gray, J. B., & Fan, G. (2008). Classification tree using target. *Computational Statistics & Data Analysis, 52*, 1362–1372.

Guyot, J., Cilas, C., & Sache, I. (2008). Influence of host resistance and phenology on South American Leaf Blight of the rubber tree with a special consideration for its temporal dynamics. *European Journal of Plant Pathology, 120*(2), 111–124.

Hallé, F., & Martin, R. (1968). Etude de la croissance rythmique chez l'hévéa (*Hevea brasiliensis* Müll.-Arg. Euphorbiacées-Crotonoïdées). *Adansonia. série 2. 8* (4), 475–503.

Holliday, P. (1969). Dispersal of conidia of *Dothidella ulei* from *Hevea brasiliensis*. *Annals of Applied Biology, 63*, 435–447.

Holliday, P. (1970). South American Leaf Blight (Microcyclus ulei) of Hevea brasiliensis. *Phytopathological Papers, 12, Farnham Royal, CAB*, 1–31.

Hu, W., Mengersen, K., McMichael, A., & Tong, S. (2008). Temperature, air pollution and total mortality during summers in Sydney, 1994–2004. *International Journal of Biometeorology, 52*, 689–696.

Junqueira, N. T. V., Lieberei, R., Kalil Filho, A. N., & Lima, M. I. P. M. (1990). Components of partial resistance in *Hevea* clones to rubber tree leaf blight, caused by *Microcyclus ulei*. *Fitopatologia Brasileira, 15*(3), 211–214.

Mishra, A. K., & Bhattacharyya, A. (2001). Epidemiology and Management of Sigatoka Disease of Banana. *Journal of Mycology and Plant Pathology, 31*(2), 156–164.

Muller, R., & Möckel, M. (2008). Logistic regression and CART in the analysis of multimarker studies. *Clinica Chimica Acta, 394*, 1–6.

Nakache, J. P., Vilain, J., & Fertil, B. (1996). Discrimination et régression par une méthode neuromimétique et par la méthode de segmentation CART: application à différentes données et comparaison des résultats. *Revue de Statistique Appliquée, XLIV*(4), 19–40.

Perrier, X., & Delvaux, B. (1991). Une méthodologie de détection et de hiérarchie des facteurs limitant la production à l'échelle régionale. Application à la culture bananière. *Fruits, 46*(3), 213–226.

Stahel, G. (1917). De Zuid-Amerikaansche Hevea-Bladziekte veroorzaakt door *Melanopsammopsis Ulei* nov. Gen. (= *Dothidella Ulei* P. Hennings). *Surinam Department of Agriculture Bulletin 34*, 111 pp.