The use of fuzzy decision trees for coffee rust warning in Brazilian crops

Marcos E. Cintra*, Carlos A. A. Meira†, Maria C. Monard*, Heloisa A. Camargo‡, and Luiz H. A. Rodrigues§

*Mathematics and Computer Science Institute - University of Sao Paulo (USP)
CEP 13.566-590, Sao Carlos - SP, Brazil. Email: {cintra,mcmonard}@icmc.usp.br
†Embrapa Agriculture Informatics - C.P. 6041, CEP 13083-970
Campinas, SP, Brazil. Email: carlos@cnptia.embrapa.br
‡Computer Science Department - Federal University of Sao Carlos - UFSCAR
PO Box 676, 13.565-905, São Carlos - SP, Brazil. Email: heloisa@dc.ufscar.br
§School of Agricultural Engineering - State University of Campinas - Unicamp
CEP 13083-875 - Campinas, SP, Brazil. Email:lique@feagri.unicamp.br

Abstract—This paper proposes the use of fuzzy decision trees for coffee rust warning, the most economically important coffee disease in the world. The models were induced using field data collected during 8 years. Using different subsets of attributes from the original data, three distinct datasets were constructed. The class attribute, representing the monthly infection rate, was used to construct six datasets according to two distinct infection rates. Induced models can be used to trigger alerts when estimated monthly disease infection rates reach one of the two thresholds. The first threshold allows applying preventive actions, whereas the second one requires a curative action. The fuzzy decision tree models were compared to the ones induced by a classic decision tree algorithm, taking into account the accuracy and the syntactic complexity of the models, as well as its quality according to an expert opinion. The fuzzy models showed better accuracy power and interpretability.

Keywords-coffee rust disease; machine learning; decision trees; fuzzy logic; fuzzy decision trees.

I. INTRODUCTION

Coffee rust (*Hemileia vastatrix* Berk. & Br.) is the main disease in coffee crops (*Coffea arabica* L.) in the world. In regions of Brazil, where climate conditions favour the disease, losses can reach about 35%, and sometimes even more than 50% [1]. In addition to its economical importance, the coffee rust disease meets other requirements that justify the development of forecast and warning models, according to its variations among crop seasons and availability of economically viable control measures.

Several empirical models to predict some features of coffee rust have been proposed. Adjusting the observed data to regression equations is the most common modeling technique [2]. In [3], a model to explain the biological action course of the *H. vastatrix* is proposed. The values obtained with this model from data observed in the field were used in the development of regression equations to predict the progress rate of the disease [4]. In [5], the potential of neural networks to tackle the coffee rust epidemic is evaluated. In [6], a prediction model using a calculated matrix based on the severity values of the coffee rust is empirically evaluated.

The induction of decision trees (DTs) is an alternative modeling technique. DTs are easy to interpret. Furthermore, the visual approach of DTs is particularly helpful in comprehending sequential decisions and outcome dependencies. Differently from regression techniques, the multicollinearity among independent variables does not affect the performance of the DTs [7]. Also, several variables, numeric and categoric, can be analyzed at the same time, since the DT induction algorithm itself takes care of selecting the most relevant variables (features).

Decision trees were used in [8] to model and predict severity categories of gray leaf spot of maize, as well as in [9] to predict the risk (high or low) of mortality of pine trees due to the Annosus root disease. Specifically related to coffee rust, epidemics of the disease were analyzed with the aid of DT [10]. Decision trees have also been used to predict the infection rate of coffee rust in [11].

Fuzzy decision trees are based on the fuzzy theory proposed by Zadeh [12], which can be considered an extension of the classic set theory. FUZZYDT, described in [13], is a fuzzy DT algorithm based on the classic C4.5 algorithm proposed by Quilan [14]. It uses the same measures of entropy and information gain as C4.5 to recursively induce the model. FUZZYDT fuzzifies the examples of the dataset according to the fuzzy definition of the attributes. This way, the process can be seen as generating a DT using categoric attributes. According to the fuzzy reasoning mechanism, several rules may have a certain compatibility with the input example and are then triggered at the same time. The classification decision can be done using the class in the consequent of the rule with highest compatibility with the input example, or the class with highest sum of compatibility with the input example.

In this paper we present the fuzzy DTs induced by FUZZYDT and compare them with the warning models generated by a classic DT algorithm. Experimental results show the superiority of the fuzzy decision tree models.

The remainder of this paper is organized as follows. Section II describes the data collection, decision trees, including

the FUZZYDT algorithm, and the experiments. Section III presents results and discussion. Section IV presents the conclusions and future work.

II. MATERIAL AND METHODS

This section describes the data collection process and the generation of the datasets used in the experiments, introduces the topic of decision trees and the FUZZYDT algorithm, and presents the experiments.

A. Data Collection

The data used in this work were collected by Japiassú et al. [15] and refer to the monthly following up of the coffee rust incidence at the Experimental Farm of the Procafé Foundation, in Varginha, Minas Gerais, Brazil (21°34'0"S, 45°24'22"W), during 8 years (October, 1998 – October, 2006). For each year, four plots with large spacing (about 4,000 plants/ha) and four dense plots (about 8,000 plants/ha) were selected. For each spacing, two plots with large fruit load (above 30 bags/ha) were selected. The sampling method adopted is recommended by Chalfoun [16]. The final dataset includes 182 examples from the total 192 available ones (monthly collected during 8 years for two different spacings). The remaining 10 samples were discarded due to problems in the collection process. No disease control was done for the agricultural year in the selected plots.

Meteorological data, such as air temperature (average, maximum, and minimum), pluvial precipitation, and relative humidity of the air, were recorded every 30 minutes by an automatic meteorological station (Davis company, Groweather Industrial model) installed close to the places of evaluation of the coffee rust incidence. The coffee rust progress between one evaluation and the following, *i.e.*, the infection rate, was defined as the dependent variable. Monthly infection rates were calculated by subtracting the disease incidence of the current month from the incidence of the previous one. The numerical values of the infection rates were mapped into two categories (or classes) [11].

The first option of the binary infection rate was done by creating the INF_RATE_G5 variable, with value 1 for infection rates equal or greater than 5 percentage points (pp) and 0 otherwise. To allow further comparisons, the INF_RATE_G10 variable was created, with value 1 for infection rates equal or greater than 10 pp, and 0 otherwise. The decision threshold set in 5 pp was determined based on the 5% incidence limit of the coffee rust recommended by Zambolim et al. [1] for the disease control. The decision threshold set to 10 pp was determined based in Kushalappa et al. [4], who proposed the risk limit of 10% of incidence to recommend the use of fungicide. It is important to notice that this value (10%) is close to the upper limit of 12% of sick leaves recommended for the use of systemic fungicides [1].

The meteorological predictive attributes were derived for probable periods of infection of the *H. vastatrix* determined by estimates of the incubation period of the fungus (calculated with a regression equation). Some of these attributes aim at contemplating known results and aspects of the disease epidemics found in the literature. Further details on the data preparation steps can be found in [17].

B. Decision Trees and FUZZYDT

A DT is a hierarchical model, which is widely used by the machine learning community, implementing the divide-and-conquer strategy. Moreover, it is an efficient non-parametric method. The well-known C4.5 DT algorithm generates a decision tree by recursively creating nodes labeled by the features selected at each step, using the information gain and entropy measures when deciding on the importance of the features. Regarding the pruning process, C4.5 uses post-pruning, *i.e.*, the pruning takes place after the tree is completely induced. The pruning process basically assesses the error rates of the tree and its components directly on the set of training examples [14]. The default confidence limits used by C4.5 is 25%. By default, leaf nodes are created when at least 2 examples fit the node.

FUZZYDT, proposed in [13], uses the same measures of C4.5 (entropy and information gain) to decide on the importance of the features. However, the features are all defined in terms of fuzzy sets, and the training set is fuzzified before the decision tree induction. This way, the process can be seen as inducing a tree using only discrete features.

Classic DTs can be seen as a set of disjunct rules in which only one rule is fired to classify a new example. Differently, fuzzy decision trees can be seen as a set of rules which can be fired simultaneously, each one with a degree of compatibility, in order to classify a new example. To this end, the classic or the general fuzzy reasoning methods can be applied. The classic fuzzy reasoning method uses the class of the rule with highest compatibility to the input example to classify it, while the general fuzzy reasoning method uses the class with highest sum of compatibility among all rules with the same class in the consequent. In other words, a classic DT classifies a new example by checking the root test and then following the next triggered branch of the tree, until reaching a leaf. FUZZYDT, on the other hand, calculates a membership degree for the input values in each fuzzy set defining the attributes. This way, it is possible to compute a confidence degree for each rule which covers the input example. Since all branches might be fired, this confidence degree is used by the classification process giving credibility to the final classification decision. Observe that for a classic DT, whenever the input values are located in the decision frontiers, misclassification might occur.

C. Experiments

Experiments were carried out with J48, a C4.5 implementation in the WEKA [18] framework, and with FUZZYDT. Our own implementation of the FUZZYDT was used. Both

algorithms were executed using default parameters, except the minimum number of examples in terminal nodes, which was set to 5. The following aspects were considered to compare the results:

- Performance, in terms of error rates;
- Interpretability of the generated models, in terms of their syntactic complexity, taking into consideration the number of rules generated and the number of conjunctions of these rules, i.e. the rule length.

Table I describes the general characteristics of the datasets used (Dataset), presenting number of features, including the number of continuous and discrete features, respectively, in brackets, number of classes, majority error (ME) – representing the error obtained by the most naive classifier, i.e. a classifier which classifies all examples using the dataset majority class –, and number of fuzzy sets defining the attributes (FS). The number of fuzzy sets was empirically defined. Triangular fuzzy sets evenly distributed in the partitions were used. The experiments with FUZZYDT used the classic fuzzy reasoning method, which classifies a new example using the class of the rule with the highest compatibility with it.

Table I: Characteristics of the datasets and number of fuzzy sets defining the attributes.

Dataset	Features	Classes	ME	FS
M1G5	23 (22, 1)	2	46.15	3
M1G10	23 (22, 1)	2	29.12	3
M2G5	14 (13, 1)	2	46.15	3
M2G10	14 (13, 1)	2	29.12	3
M3G5	11 (10, 1)	2	46.15	3
M3G10	11 (10, 1)	2	29.12	3

The names of the datasets are intended to present information regarding the set of variables used and the infection rate for the class. Three predictive attribute selections were chosen. The first option (M1) included all 23 predictive attributes. Option M2 included attributes derived from meteorological variables measured by sensors available in most meteorological stations. This way, we aim at widening the use of the models. M3 excluded from M2 the attributes whose preparation depended on hour recordings. This is due to the fact that, in general, available public data from meteorological stations of other institutions is often daily summarized. Again, the motivation here was to widen the use of the induced models. Both, the attributes and the three dataset options (M1, M2 and M3) are listed in Table II. G5 and G10 refer to the infection rates, *i.e.*, the binary class variable, separating examples according to the infection rate thresholds of 5 pp and 10 pp, respectively.

Table II describes the variable names, where the two first ones are class related, type (T: binary (B) or numeric (N)), unit (U) (mm: millimetres; h: hour; mm/h: millimetres per hour; km/h: kilometres per hour; °C: degrees Celsius; %: percentage), and descriptions of all attributes. Columns named 1, 2, and 3 present a check mark for the variables included in the corresponding datasets M1, M2, and M3.

Table II: Description of the variables.

Variable Name	T	U	Description	1	2	3
INF_RATE_G5	B	-	Infection rate 1 if ≥ 5 pp, 0 otherwise	✓	✓	✓
INF_RATE_G10	B	-	Infection rate 1 if ≥ 10 pp, 0 otherwise	✓	✓	✓
01 - RAINY_DAYS	N	days	Number of rainy days (precipitation ≥ 1 mm)	✓	✓	✓
02 - SPACING	B	-	Spacing: dense or wide crops	✓	✓	✓
03 - RAIN_PREC_AVG	N	mm	Average of the daily rain precipitation	✓		
04 - AVG_MAX_RAIN	N	mm/h	Average of the maximum daily rain precipitation	✓	✓	✓
05 - NH_RH_95	N	h	Daily average of night hours with air relative humidity ≥ 95%	✓	✓	
06 - DH_RH_95	N	h	Daily average of hours when air relative humidity ≥ 95%	✓	✓	
07 - AC_RAIN_PREC	N	mm	Rain precipitation accumulated	✓	✓	✓
08 - SUM_NH_RH_95	N	h	Sum of NH_RH_95	✓		
09 - SUM_DH_RH_95	N	h	Sum of DH_RH_95	✓		
10 - WIND_SUM_AVG	N	km/h	Wind speed sum average	✓		
11 - D_TEMP_AVG_95	N	°C	Average daily temperature when air relative humidity ≥ 95%	✓	✓	
12 - AVG_MAX_TEMP	N	°C	Average of the maximum daily temperatures	✓	✓	✓
13 - AVG_MAX_T_IP	N	°C	Average of the maximum daily temperatures for incubation period	✓	✓	✓
14 - AVG_TEMP	N	°C	Average of the temperatures	✓	✓	✓
15 - AVG_T_IP	N	°C	Average of the daily temperatures for incubation period	✓	✓	✓
16 - AVG_MIN_TEMP	N	°C	Average of the minimum daily temperatures	✓	✓	✓
17 - AVG_MIN_T_IP	N	°C	Average of the minimum daily temperatures for incubation period	✓	✓	✓
18 - REL_HUM	N	%	Daily air relative humidity	✓	✓	✓
19 - WIND_AVG	N	km/h	Average of the daily wind speed	✓		
20 - ACC_D_INFEC	N	-	Daily accumulated infection value	✓		
21 - UID	N	days	Number of unfavourable days for infection	✓		
22 - FVFID	N	days	Number of favourable and very favourable days for infection	✓		
23 - VFID	N	days	Number of very favourable days for infection	✓		

III. RESULTS AND DISCUSSION

The focus of the coffee rust control must be placed on large fruit load years, when the disease progress is faster and the attack more severe. The warning models presented in this section predict the infection of the following month based on 5 pp and 10 pp increase rates. These models might be helpful to decide on when and what type of measures must be taken for the disease control.

Both methods presented similar execution time. Regarding the accuracy of the induced models, Table III presents the majority error of each dataset (ME) and the error rates (Error) and standard deviation (SD) obtained by the FUZZYDT and J48 algorithms. The error rates were calculated using the 10-fold cross-validation strategy. The smaller error rate for each model is light-grey shaded.

It can be observed that FUZZYDT obtained the smallest error rates for all datasets. However, for both algorithms, J48 and FUZZYDT, the error rate standard deviations were considerably high.

Table III: Error rates

Dataset	ME	FuzzyDT Error	SD	J48 Error	SD
M1G5	46.15	15.29	9.56	18.72	8.59
M1G10	29.12	17.06	8.50	21.86	8.23
M2G5	46.15	16.47	12.83	17.62	7.91
M2G10	29.12	17.06	9.65	23.53	10.08
M3G5	46.15	16.47	13.62	18.07	8.71
M3G10	29.12	16.47	9.77	21.36	9.59

Aiming to verify if the high error rate standard deviations were due to the kind of model induced (DTs), we investigated the results obtained by other kinds of models. To this end, the following learning algorithms, all of them available in the WEKA framework, were executed with default parameters: Nearest Neighbor using 3 neighbors (which classifies an input with the class of the closest examples to it); Naive Bayes (a probabilistic model); Random Forests (which induces rules based on the divide and conquer strategy and partial decision trees); and a multi-layer perceptron (a neural network approach). For all these learning algorithms the error rates were higher than those obtained by FuzzyDT, having also high standard deviations. This shows that the problem is not related to the kind of model induced. For the DT models, the high error rate standard deviations indicate that there are some rules with high misclassification rates. This problem can be tackled by focusing on these rules in order to search for possible exceptions, as proposed in [19]. Moreover, for FuzzyDT, the definition of the class as a fuzzy attribute may also help to alleviate the problem.

To test whether there was a statistically significant difference among the algorithms, the Wilcoxon [20] matched pair test and the Mann-Whitney [21] test were performed. Both tests stated that FuzzyDT and J48 differ significantly with 95% confidence.

Table IV presents the number of rules (Rules), total number of conjunctions (Conjs), and the number of attributes present in each model (Attribs) for the FuzzyDT and J48 algorithms, of the models generated with all the training examples. The smaller number of rules, conjunctions, and attributes for each model are indicated by a light-grey shade.

Table IV: Number of rules, conjunctions, and attributes.

Dataset	FuzzyDT Rules	Conjs	Attribs	J48 Rules	Conjs	Attribs
M1G5	5	8	2	6	20	4
M1G10	7	15	3	8	34	7
M2G5	3	3	1	5	13	3
M2G10	9	22	4	4	9	3
M3G5	7	15	3	5	12	4
M3G10	7	15	3	4	9	3

Regarding the number of rules and conjunctions, FuzzyDT presented smaller values for the first three datasets, while J48 for the remaining ones. It is interesting to notice that for the first three datasets, although the difference

in the number of rules is not expressive, J48 presented a considerable larger number of conjunctions when compared to FuzzyDT. In other words, J48 induced more specialized rules. On the other hand, FuzzyDT presented a larger number of conjunctions for the M2G10 dataset.

Another important issue regarding the interpretability of the models is the number of attributes they include. FuzzyDT presented a smaller number of attributes for four datasets and a tie for one dataset (M3G10), while J48 presented less attributes for only one dataset (M2G10). Notice that although FuzzyDT produced larger numbers of rules and conjunctions for the last three datasets, the number of attributes included in the models is not larger than the ones included in the J48 induced models. In fact, the number of fuzzy sets is directly related to the number of rules produced in the final models.

For the M1G5 dataset, the model induced by J48 presented an issue related to classic decision trees: the repeated use of the same numerical attribute in a single branch/rule. This is shown in the following rules/branches (underlined). Note that these two rules represent only part of the induced model. This characteristic makes it harder to understand the whole model.

1) If UID is ≤ 25 and AVG_MAX_T_IP is ≤ 28.6 and D_TEMP_AVG_95 is ≤ 17.6 and WIND_SUM_AVG is ≤ 2.9 and D_TEMP_AVG_95 is ≤ 16.3 then class is 0
2) If UID is ≤ 25 and AVG_MAX_T_IP is ≤ 28.6 and D_TEMP_AVG_95 is ≤ 17.6 and WIND_SUM_AVG is ≤ 2.9 and D_TEMP_AVG_95 is > 16.3 then class is 1

In order to allow further analysis, next we present and discuss the model induced by J48 for the M3G10 dataset — Figure 1. This model uses three attributes with binary outcome tests. According to this model, the effect of the daily air relative humidity (REL_HUM) has prevailed if compared to the temperature influence. REL_HUM is included with higher importance in the faster evolution periods of the coffee rust. Higher average temperatures in the incubation period (AVG_MAX_T_IP) and lower daily average temperatures (AVG_TEMP) during the infection period had a negative effect on the infection rates, confirming epidemiologic studies of the disease [10].

For the fuzzy model, first we present it as a set of rules, followed by the tree in Figure 2.

- IF AVG_MIN_TEMP is LOW then CLASS is 0
- IF AVG_MIN_TEMP is MEDIUM then CLASS is 0
- IF AVG_MIN_TEMP is HIGH and REL_HUM is LOW then CLASS is 1
- IF AVG_MIN_TEMP is HIGH and REL_HUM is MEDIUM then CLASS is 0
- IF AVG_MIN_TEMP is HIGH and REL_HUM is HIGH and AVG_MAX_T_IP is LOW then CLASS is 1
- IF AVG_MIN_TEMP is HIGH and REL_HUM is HIGH and AVG_MAX_T_IP is MEDIUM then CLASS is 1
- IF AVG_MIN_TEMP is HIGH and REL_HUM is HIGH and AVG_MAX_T_IP is HIGH then CLASS is 0

Although the model also uses three attributes, one of them differs from the J48 model (for the FuzzyDT model, the average minimum temperature is used, while J48 uses
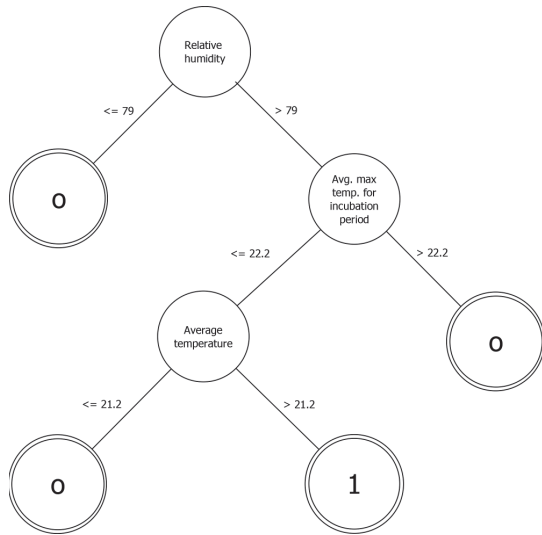
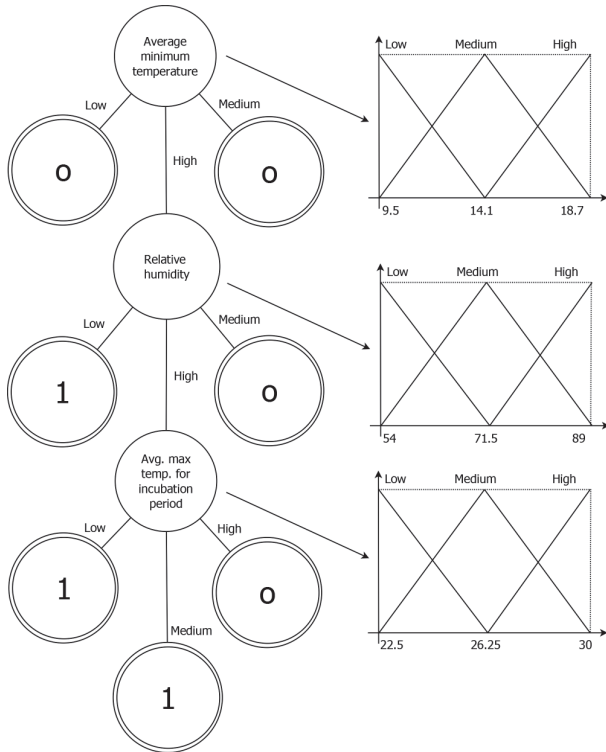Figure 1: Decision tree generated by J48 for M3G10.

Figure 2: Decision tree generated by FUZZYDT for M3G10.

the average temperature). From the interpretability point of view, it is more pleasant to read the fuzzy model. Although the linguistic values must be representative for the user, the fuzzy model does not have unnatural separations of attributes as in the J48 model, which complicates the understanding of the model and force different classifications for very similar examples. To illustrate this issue, consider two input examples, with 80 as relative humidity, 22.2 for

the average maximum temperature for the incubation period, but different average temperatures, the first having 21.2 and the second 21.3; although they are almost identical, the J48 model will classify the first input as class 0, while the second one will be classified as class 1. In cases like this, it is difficult to justify the classifications. The fuzzy model, on the other hand, will use the compatibility degrees of each rule (or branch) to justify the classification.

The fuzzy model shows the importance of the daily minimum temperatures average attribute (AVG_MIN_TEMP). The leaf wetness necessary for the germination of the spores of the *H. vastatrix* generally happens during the night period, when the minimum daily temperature also occurs. The other two variables are the same as the J48 model, with compatible behaviors in both models, with exception to the rule "IF AVG_MIN_TEMP is HIGH and REL_HUM is LOW THEN CLASS is 1". From the expert point of view, specifically for this rule, class 0 would be expected.

For large fruit load years, late atomizations are not recommended after the confirmation that the incidence levels of the coffee rust is higher than 5% [2]. This way, the MxG5 models are more indicated for the decision support of the disease control. MxG10 models, on the other hand, might be used as additional instruments informing and alerting about the fact that measures taken should be more urgent and/or effective, based on the fact that conditions are favourable for an even more accelerated development of the disease.

IV. CONCLUSION

The coffee rust disease can be found in all Brazilian coffee crops. It is the most important coffee disease in the world. The Brazilian losses in regions with favorable conditions to the disease reach, on average, about 35%. Since this disease has a high economical impact, several models have been proposed for its control and warning.

In this paper, we presented and compared two decision tree methods for the warning of the coffee rust disease: a fuzzy and a classic model. Data collected in the field were used to form six datasets based on two different infection rate levels (5 pp and 10 pp) and three distinct subsets of selected attributes (M1, M2, and M3). Fuzzy and classic decision trees were induced for all datasets.

The classic model used in this paper is the J48, an implementation of C4.5. The fuzzy model, namely FUZZYDT, is based on the classic C4.5 decision tree, but incorporates all the interesting characteristics of the fuzzy logic related to interpretability and handling of continuous attributes.

FUZZYDT presented competitive error rates and models. FUZZYDT also has the advantage of producing better interpretable models with no unnatural discretization of the attributes. It is important to highlight the fact that, although the fuzzy models showed better performance for all datasets, the improvement, when compared to the J48 models, was more expressive for the MxG10 datasets.

Regarding the selection of features for the definition of the three models (M1, M2, and M3), models M3 (with the smallest subset of attributes) had competitive performance, having the advantage that meteorological hourly records are not included.

As future work, we intend to tackle the high standard deviation rates by searching for exception rules to be included in the final classification model. We also intend to transform the class attribute into a fuzzy variable, considering three linguistic labels related to the infection rate: $low \leq 5$ pp, 5 pp $< medium \leq 10$ pp, and $high > 10$ pp.

REFERENCES

[1] L. Zambolim, F. X. R. Vale, A. A. Pereira, and G. M. Chaves, *Coffee (Coffea arabica L.): diseases control - diseases caused by fungi, bacterium, viruses. In: Plant diseases control: large cultures.*, F. X. R. Vale and L. Zambolim, Eds. Federal University of Varginha - UFV, 1997, vol. 1, (in Portuguese).

[2] L. Zambolim, F. X. R. Vale, H. Costa, A. A. Pereira, and G. M. Chaves, *The state-of-the-art of the coffee production technologies. Epidemiology and integrated control of the coffee rust*, L. Zambolim, Ed. Suprema, 2002, (in Portuguese).

[3] A. C. Kushalappa, M. Akutsu, and A. Ludwig, "Application of survival ratio for monocyclic process of *Hemileia vastatrix* in predicting coffee rust infection rates," *Phytopathology*, vol. 73, pp. 96–103, 1983.

[4] A. C. Kushalappa, M. Akutsu, S. H. Oseguera, G. M. Chaves, and C. Melles, "Equations for predicting the rate of coffee rust development based on net survival ratio for monocyclic process of *Hemileia vastatrix*," *Fitopatologia Brasileira*, vol. 9, pp. 255–271, 1984.

[5] A. C. S. Pinto, E. A. Pozza, P. E. Souza, A. A. A. Pozza, V. Talamini, J. M. Boldini, and F. S. Santos, "Description of epidemics of coffee rust with neural network," *Fitopatologia Brasileira*, vol. 27, pp. 517–524, 2002, (in Portuguese).

[6] C. L. P. Garcon, L. Zambolim, E. S. G. Mizubuti, F. X. R. Vale, and H. Costa, "Coffee leaf rust control based on rust severity values," *Fitopatologia Brasileira*, vol. 29, pp. 486–491, 2004, (in Portuguese).

[7] D. J. Butt and D. J. Royle, *Multiple regression analysis in the epidemiology of plant diseases. In: Epidemics of plant diseases: mathematical analysis and modeling*, 2nd ed., J. Kranz, Ed. Springer-Verlag, 1990.

[8] P. A. Paul and G. P. Munkvold, "A model-based approach to preplanting risk assessment for gray leaf spot of maize," *Phytopathology*, vol. 94, pp. 1350–1357, 2004.

[9] F. A. Baker, D. L. Verbyla, C. S. Hodges, and E. W. Ross, "Classification and regression tree analysis for assessing hazard of pine mortality caused by *Heterobasidion annosum*." *Plant Disease*, vol. 77, pp. 136–139, 1993.

[10] C. A. A. Meira, L. H. A. Rodrigues, and S. A. Moraes, "Analysis of coffee leaf rust epidemics with decision tree," *Tropical Plant Pathology*, vol. 33, pp. 114–124, 2008, (in Portuguese).

[11] C. A. A. Meira, L. H. A. Rodrigues, and S. A. Moraes, "Warning models for coffee rust control in growing areas with large fruit load," *Pesquisa Agropecuaria Brasileira*, vol. 44, pp. 233–242, 2009, (in Portuguese).

[12] L. Zadeh, "Fuzzy sets," *Information and Control*, vol. 8, pp. 338–353, 1965.

[13] M. E. Cintra, M. C. Monard, and H. A. Camargo, "An evaluation of rule-based classification models induced by a fuzzy method and two classic learning algorithms," *IEEE - The Brazilian Symposium on Artificial Neural network (SBRN)*, vol. 1, pp. 188–193, 2010.

[14] J. R. Quinlan, *C4.5: Programs for Machine Learning (Morgan Kaufmann Series in Machine Learning)*, 1st ed. Morgan Kaufmann, January 1993.

[15] L. B. Japiassu, A. W. R. Garcia, A. E. Miguel, C. H. S. Carvalho, R. A. Ferreira, L. Padilha, and J. B. Matiello, "Effect of crop load, tree density and weather conditions on the development of the coffee leaf rust." *Simpósio de pesquisa dos cafés do Brasil*, 2007, 5° edition.

[16] S. M. Chalfoun, *Coffee tree diseases: importance, identification and control methods*. Federal University of Lavras, Brazil, FAEPE (Foundation for the Support of Teaching, Researching and Extension), 1997, (in Portuguese).

[17] C. A. A. Meira, "Process of knowledge discovery in databases for analysis and warning of crop diseases and its application on coffee rust," Ph.D. dissertation, State University of Campinas, Brazil, 2008, (in Portuguese).

[18] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*, 3rd ed. Morgan Kaufmann, 2011.

[19] R. C. Prati, M. C. Monard, and A. C. P. L. F. Carvallho, "Looking for exceptions on knowledge rules induced from HIV cleavage data set," *Genetics and Molecular Biology*, vol. 27, no. 4, pp. 637–643, 2004.

[20] J. Demsar, "Statistical comparison of classifiers over multiple data sets," *Journal of Machine Learning Research*, vol. 7, no. 1, pp. 1–30, 2006.

[21] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrics Bulletin*, vol. 1, pp. 80–83, 1945.