Lack of Data: Is It Enough Estimating the Coffee Rust with Meteorological Time Series?

David Camilo Corrales[1,2](✉), German Gutierrez[2], Jhonn Pablo Rodriguez[1](✉), Agapito Ledezma[2], and Juan Carlos Corrales[1](✉)

[1] Telematics Engineering Group, University of Cauca, Popayán, Colombia
{dcorrales,jhonnpablo,jcorral}@unicauca.edu.co

[2] Computer Science Department, Carlos III University of Madrid, Leganes, Spain
davidcamilo.corrales@alumnos.uc3m.es
{ggutierr,ledezma}@inf.uc3m.es
http://www.unicauca.edu.co,http://www.uc3m.es

Abstract. Rust is the most economically important coffee disease in the world. Coffee rust epidemics have affected a number of countries including: Colombia, Brazil and Central America. Researchers try to predict the Incidence Rate of Rust (IRR) through supervised learning models, nevertheless the available IRR measurements are few, then the data set does not represent a sample trustworthy of the population. In this paper we use Cubic Spline Interpolation algorithm to increase the measurements of Incidence Rate of Rust and subsequently we construct different subsets of meteorological time series: (i) *Daily meteorology*, (ii) *Meteorological variation*, and (iii) *Previous meteorology* using M5 Regression Tree, Support Vector Regression and Multi-Layer Perceptron. *Previous meteorology* with Multi-Layer Perceptron have shown better results in measures as Pearson Coefficient Correlation of 0.81 and Mean Absolute Error = 7.41%.

Keywords: Coffee rust · Incidence rate of rust · Regression models · Time series · Interpolation

1 Introduction

Coffee rust is the main disease in coffee crops in the world [5]. The Rust disease has reduced considerably the coffee production in Colombia (by 31% on average during the epidemic years compared with 2007), Brazil (where climate conditions favour the disease, losses can reach about 35%, and sometimes even more than 50%) [5] and Central America (by 16% in 2013 compared with 2011–2012 and by 10% in 2013–2014 compared with 2012–2013). These reductions have had direct impacts on the livelihoods of thousands of smallholders and harvesters. For these populations, particularly in Central America, coffee is often the only source of income used to buy food and supplies for the cultivation of basic grains. As a result, the coffee rust epidemic has had indirect impacts on food security [2].

O. Gervasi et al. (Eds.): ICCSA 2017, Part II, LNCS 10405, pp. 3–16, 2017.
DOI: 10.1007/978-3-319-62395-5_1

Several supervised learning models to predict some features of coffee rust have been proposed. Brazilian researchers monitoring the coffee rust incidence through Fuzzy Decision Trees, Bayesian Networks, Support Vector Machine at the Experimental Farm of the Procafé Foundation, in Varginha, Minas Gerais, Brazil, during 8 years (October, 1998–October, 2006), which includes 182 examples and 23 attributes that involves weather conditions and physical properties of crop.

In [17] the authors use SVM regression; rather than predicting the change in incidence, they consider actual incidence. They trained a SVM that provided a correlation of about 0.94 between predicted and actual incidences. However, when they try to devise an alarm system for predicting values above a given threshold, the number of false negatives turned to be too high.

Support Vector Regression was adapted interpreting intervals as classes in [16]. This approach presents a framework where the costs of false negatives are higher than that of false positives, and both are higher than the cost of warning predictions.

In [5] the authors compare two decision tree methods for the warning of the coffee rust disease: a fuzzy and a classic model. They built six datasets based on two different infection rate of rust levels: 5 and 10 % points. The fuzzy model, namely FUZZYDT, is based on the classical C4.5 decision tree, but incorporates all the interesting characteristics of fuzzy logic related to interpretability and handling of continuous attributes.

The approach represented in [20] constructs two Bayesian Networks (ignoring the weather variables) using the CaMML (Causal Minimum Message Length). They run CaMML with different structural priors, and tries to learn the best causal structure to take into account the data, i.e., to describe the information about relationships between variables.

Similarly, Colombian researchers use Support Vector Machine, Decision Trees, Artificial Neural Networks, Bayesian Networks, and Ensemble Methods for estimating the incidence rate of rust. The coffee rust records were collected trimonthly for 18 plots, closest to weather station at the Experimental Farm of the Supracafé Enterprise, in Cajibio, Cauca, Colombia, during the period of 2011–2013. The dataset includes 147 instances and 21 attributes that involves weather conditions, physical properties of crop, and crop management [6].

The work [9] builds three dataset based on different subset of attributes. The dataset were evaluated with three classifiers: Support Vector Regression, Backpropagation Neural Network, and M5 Regression Tree. These classifiers provide a correlation of about 0.47, 0.4549, and 0.4532 to the dataset with 13 attributes.

In [10] it is proposed theoretically an early warning system for coffee rust based on Error Correcting Output Codes with binary Support Vector Machines. The authors suggest to build a dataset with features as: plant density, shadow level, soil acidity, previous month rainfall, previous nighttime rainfall intensity, and relative humidity in the previous days.

The authors [15] propose a rule extraction approach to detect coffee rust from decision tree induction and expert knowledge using graph-based representation. The patterns found were modeled according to meteorological variables related to coffee rust dataset.

The works [7,8] develop an empirical multi-classifier based on Cascade Generalization method. The multi-classifier is composed by: backpropagation neural network to select among two experts classifiers: Regression tree (M5) to detect the incidence rate of rust greater than 7.18%; or support vector regression to detect the incidence rate of rust less than 7.18%.

However, the main drawback of the related works mentioned above is the low number of instances to try to predict a continuous value: Incidence Rate of Rust (values are 0%–100%); if the available examples are few, the data set does not represent a sample trustworthy of the population, then the classifiers will be not inaccurate [7]. This paper proposes three data set for estimating Incidence Rate of Rust through meteorological time series. We used cubic spline interpolation to increase the measurements of Incidence Rate of Rust. The remainder of this paper is organized as follows: Sect. 2 describes the coffee rust context and the regression models used; Sect. 3 the three data set proposed for estimating Incidence Rate of Rust through meteorological time series; Sect. 4 presents experimental results and Sect. 5 conclusions and future works.

2 Background

In this section we explained the coffee rust context, the dataset employed, cubic spline interpolation algorithm, concepts of times series, and the regression models used.

2.1 Coffee Rust

Coffee rust is caused by the fungus *Hemileia vastatrix*, a parasite that affects the leaves of the genus *Coffea*. Among the cultivated species, *Coffea arabica* is the most severely attacked. The disease causes defoliation that, when acute, can lead to the death of branches and heavy crop losses. The first symptoms are small yellowish spots that appear on the underside of leaves, where the fungus has penetrated through the stomata. These spots then grow, coalesce and produce uredospores with their distinctive orange colour. Chlorotic spots can be observed on the upper surface of the leaves. During the last stage of the disease, the spots become necrotic [2]. The progression of coffee rust depends on four factors that appears simultaneously [23]:

- **The host:** variety of coffee plants which are susceptible and resistant to rust. Varieties as *Típica*, *Borbón* and *Caturra* suffer severe rust attacks, while *Castillo* multiline variety is highly resistant to rust.
- **Pathogenic organism:** *Hemileia vastatrix* lifecycle begins with the germination of uredospores in 2–4 h under optimum conditions. Within 24–48 h, infection is completed. Once the infection is completed, the underside of the leaf is colonized and sporulation will occur through the stomata [19].

- **Weather conditions:** constant precipitations mainly in the afternoon and night with cloudy sky, high humidity in the plants and low temperatures are relevant factors for germination of rust. Disease spread and development is usually limited to the rainy season, while in dry periods the rust incidence is very low.
- **Agronomic practices:** is referred to properties of crop sowing (plant spacing, percentage of the shade, etc.), application of fungicides and fertilizers on coffee crops with aim to avoid several rust attacks.

In Colombia the rust progress is evaluated through the methodology developed by Centro Nacional de Investigaciones de Café (Cenicafé) [23], which is explained as follows:

Measurement Process of Rust in Colombian Coffee Crops: The Incidence Rate of Rust (IRR) is calculated for a plot with an area lower or equal to one hectare. The methodology is composed of three steps: (i) Farmer must be standing in the middle of the first furrow and he has to choose one coffee tree and pick out the branch with greater foliage for each level (high, medium, low); the leaves of the selected branches are counted as well as the infected ones with rust. (ii) Farmer must repeat step (i) for every tree in the plot until 60 trees are selected. The farmer must take in consideration that the same number of trees must be selected in every furrow (e.g. if plot has 30 furrows, the farmer selects two coffee trees for each furrow). (iii) Finalized step (i) and (ii), the leaves of the coffee trees selected (LCT) are added as well as the infected leaves with rust (ILR). Later it must be computed the Incidence Rate of Rust (IRR) using the following formula:

$$IRR = \frac{ILR}{LCT}\,100 \tag{1}$$

For the purposes of this research, IRR was used as the dependent variable, collected by Cenicafé in the next Colombian region:

Jazmín Village: The data used in this paper was obtained from Jazmín Village which is a coffee growing area sowing with *Caturra* variety in 45 farms approximately, monitored by Cenicafé and located in Santa Rosa de Cabal, Colombia ($4°55'00''$N, $75°38'0''$W). The initial data set for this study records 43 measurements of Incidence Rate of Rust (average of 30 coffee trees per measurement) and 1024 samples for six daily meteorological attributes around 26/02/1986 and 15/12/1988. The input features available were: minimum, average and maximum temperature (*MinTem, AvgTem, MaxTem*), sun hours (*SunHours*), accumulated rain during day and night (*RainDay, RainNight*).

2.2 Time Series

A time series is just a sequence of time values which are usually regularly sampled in time: $(x_1, x_2, ..., x_{t-1}, x_t)$. There are univariate time series and multivariate time series. Univariate time series refers to a sequence of one scalar measure (e.g. Temperature at main cities taken hourly or twice a day). Additionally, multivariate time series show several different measures (or time series variables), and the aim of the studies related with them is *"to model and explain the interactions and co-movements among a group of time series variables"* [27]. Time series are used in regression models to try to give an answer for the relationship between its past and future values (forecast). Below we present the regression models used.

2.3 Regression Models

In this section we briefly describe the different regression models applied to forecast the IRR. In this work we have applied three techniques to achieve a regression model, one of them a linear model, M5 regression tree, and two non-linear models Support Vector Regression and Multilayer Perceptron, which are also universal approximators [14].

M5 Regression Tree, developed by Quinlan [21], and its result is a piecewise linear model based on trees. Some of the characteristic that make M5 more suitable than other algorithms based on regression trees (CART, classification and regression trees [3]; MARS, Multivariate Adaptive Regression Splines [11]) are the following: the computational requirements of M5 grows in a way that can tackle problems with high dimensionality, and the trees generated with M5 are smaller than other methods as CART and MARS.

Support Vector Regression (SVR): A Support Vector Machine tries to get a hyperplane as a boundary for decision so that the separation between the patterns at each side (one side for each class) of the hyperplane is maximum. A hyperplane is a subspace which number of dimensions is just one less that surrounds the hyperplane (e.g. a line is an hyperplane within a surface, e.g. a plane is a hyperplane in a 3-Dimensional space. In fact a hyperplane split its ambient space in just two parts. The approach that SVM follows to define the hyperplanes is based on statistical learning theory (see [26] for a better understanding of the topic), so that SVM is an approximate implementation of Structural Risk Minimisation (SRM) principle, instead of ANN that employs Empirical Risk Minimisation (ERM) principle ("SRM minimises an upper bound on the expected risk, as opposed to ERM that minimises the error on the training data" [12]). SVM can be also applied to regression tasks [25], in this case a distance measure is included into the loss function [24].

Multi-Layer Perceptron: Is a feed-forward artificial neural network [14], which is a universal approximator. The MLP is made up by nodes that are organised in layers: input, hidden and output layers. In this MLP there are no short-cuts, so each node within a layer is connected to every node in the following layer.

The topology of the MLP (i.e. number of hidden layers, number of nodes in each hidden layer, the activation function of each neuron, etc.) is fixed, and established by an expert or follows default parameters given by a software platform. Nonetheless, the memory and the nonlinear model itself is given by the weights of each connection, which are established throughout a learning process (back-propagation learning algorithm).

3 Data Pre-processing

To increase the measurements of the Incidence Rate of Rust, we use an interpolation algorithm from 26/02/1986 and 15/12/1988. Subsequently we construct three data set to estimate the IRR based on meteorological time series, these are presented:

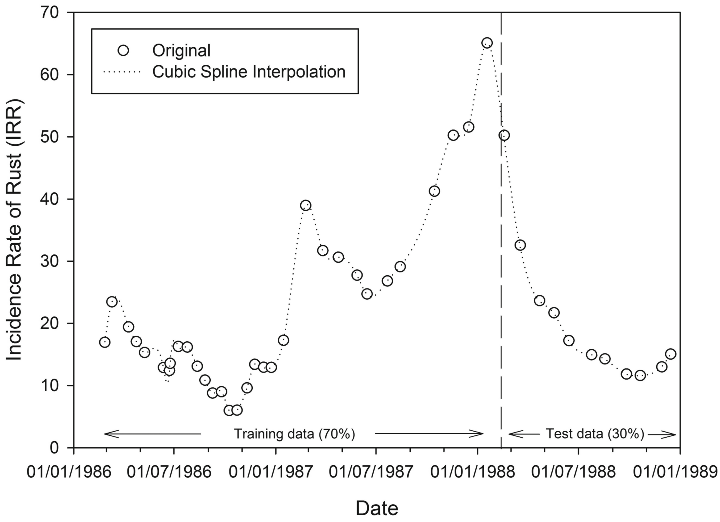


Fig. 1. Interpolation of IRR

Cubic Spline Interpolation: Having an account of the nature of problem, we decide to use cubic spline interpolation because the imputed data are inside the natural range of values, except in particular cases where quite extreme values are obtained [4].

The concept of spline is originated from conscripting technique of using a flexible strip known as spline to draw the smooth curve through a set of points. Interpolation task is just to approximate a function to a set (or sequence) of points with a certain precision which depends on the candidate function and on one or more parameters [18]. There are some issues about interpolation, but in this work we used spline cubic interpolation, which approximation candidate is a piecewise third-order polynomial function (so first and second derivatives of the candidate function are continuous at the nodes, i.e. the known points to be interpolated).

Figure 1 shows 43 original measurements of IRR represented by white dots. The remaining 981 black dots are interpolated data obtained by the Cubic Spline algorithm (executed with the function `splinefun` from the `stats` package of CRAN R [22]).

Once interpolated the IRR, we proposed three data set to estimate the IRR in Jazmín Village based on meteorological time series, we show the scheme below:

Daily Meteorology: Values of meteorological attributes and IRR dependent variable are organized by day:

$$IRR_t = \Big\{MinTem_t, AvgTem_t, MaxTem_t, SunHours_t, \\ RainDay_t, RainNight_t\Big\} \tag{2}$$

In (2) *IRR*, *MinTem*, *AvgTem*, *MaxTem*, *SunHours*, *RainDay*, and *RainNight* were measured at t moment, where t correspond to one day of 26/02/1986 and 15/12/1988.

Meteorological Variation (MV): Values of meteorological attributes is a variation Δt for IRR in t:

$$IRR_t = \Big\{MinTem_{\Delta t}, AvgTem_{\Delta t}, MaxTem_{\Delta t}, SunHours_{\Delta t}, \\ RainDay_{\Delta t}, RainNight_{\Delta t}\Big\} \tag{3}$$

where Δt is the absolute value of the difference of a meteorological attribute between $t-7$ and $t-1$. For instance:

$$MinTem_{\Delta t} = |MinTem_{t-7} - MinTem_{t-1}| \tag{4}$$

and Δt is applied for all meteorological attributes.

Previous Meteorology (PM): represents the last $t-7$ measurements of the meteorological attributes for current IRR:

$$IRR_t = \Big\{\Big(MinTem_k\Big)_{k=t-1}^{t-7}, \Big(AvgTem_k\Big)_{k=t-1}^{t-7}, \Big(MaxTem_k\Big)_{k=t-1}^{t-7}, \\ \Big(SunHours_k\Big)_{k=t-1}^{t-7}, \Big(RainDay_k\Big)_{k=t-1}^{t-7}, \Big(RainNight_k\Big)_{k=t-1}^{t-7}\Big\} \tag{5}$$

where each meteorological attribute is separated in seven sub-attributes corresponding to last $t-7$ measurements. For example:

$$\Big(AvgTem_k\Big)_{k=t-1}^{t-7} = \Big\{AvgTem_{t-1}, AvgTem_{t-2}, AvgTem_{t-3}, \\ AvgTem_{t-4}, AvgTem_{t-5}, AvgTem_{t-6}, AvgTem_{t-7}\Big\} \tag{6}$$

We take the last $t-7$ meteorological measurements to observe the evolution of disease considering that infection is completed at 24–48 h once the uredospores germinates.

4 Experimental Results

This section reports a number of experiments carried out to estimate the IRR in Jazmín Village presented above. The Software tool to carry out all the systems indicated in Subsect. 2.3 (M5, SVR and MLP) is the data mining software "workbench" called WEKA (Waikato Environment for Knowledge Analysis) [13], specifically version 3.6.12. In order to perform M5, SVR and MLP models and their algorithms for learning process, we have set the parameters suggested by Weka[1, 2, 3]. For instance, for MLP some of the default parameters are: 1 hidden layer with hn hidden nodes where $hn = (nInputs + nClasses)/2$, $nInputs = 6 \times 7$ (for the 6 meteorological measures for the previous 7 days) and $nClasses = 1$ as the MLP afford a regression problem.

We use Leave-p-out cross-validation with $p = 30\%$. Figures 2, 3 and 4 show the results for $p = 30\%$ of M5 Regression Tree, Support Vector Regression and Multi-Layer Perceptron for time series proposed: (i) *Daily meteorology*, (ii) *Meteorological variation*, and (iii) *Previous meteorology*.

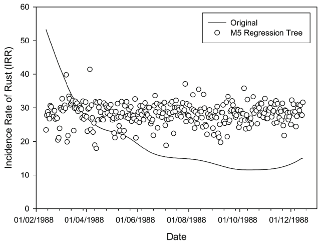
To estimate the IRR through *Daily meteorology* and *Meteorological variation* (Figs. 2 and 3 respectively), the predicted IRR are vastly different from real IRR values. Whereas the regression models of *Previous meteorology* (Fig. 4) present a similar behavior of IRR measurements of the training set but remains different of real IRR values of the test set. Multi-Layer Perceptron of *Previous meteorology* is the best approximation although some IRR predicted are outliers with values higher than 100% and less than 0%.

Table 1 shows Pearson's Correlation Coefficient (PCC) and Mean Absolute Error (MAE) to test M5 Regression Tree, Support Vector Regression and Multi-Layer Perceptron for mentioned time series. Again, the best values of the evaluation measures for all regression models were obtained for *Previous meteorology (PM)*. Multi-Layer Perceptron has the highest value of correlation (PCC = 0.81) and the least difference among predicted and real IRR measurement (MAE = 7.41%), overcoming M5 and SVR.

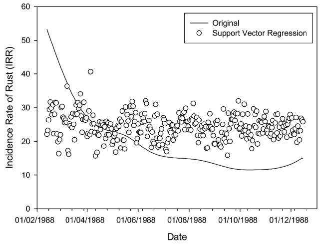
As we could see in Fig. 1 that IRR measurements of training set present periodic fluctuations among 26/02/1986–09/05/1986, 23/06/1986–09/10/1986, 23/10/1986–15/06/1987 and 16/06/1987–11/02/1988, unlike the IRR measures of test set without trend among dates 12/02/1988–15/12/1988. Hence, the regression models cannot estimate IRR values that they have not learned.

To improve the results of regression models we need more information about properties of coffee plots to analyze the seasonality behavior of IRR. We expect

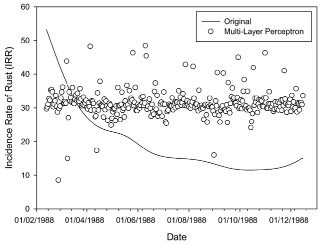
[1] http://weka.sourceforge.net/doc.dev/weka/classifiers/functions/MultilayerPerceptron.html.
[2] http://weka.sourceforge.net/doc.dev/weka/classifiers/functions/SMO.html.
[3] http://weka.sourceforge.net/doc.dev/weka/classifiers/trees/m5/M5Base.html.



M5 Regression Tree

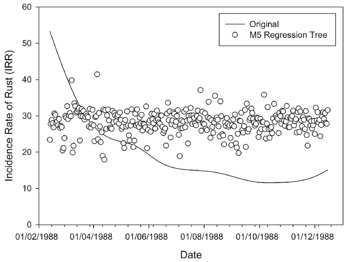


Support Vector Regression

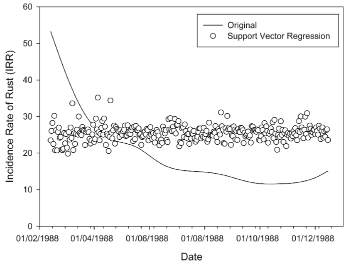


Multi-Layer Perceptron

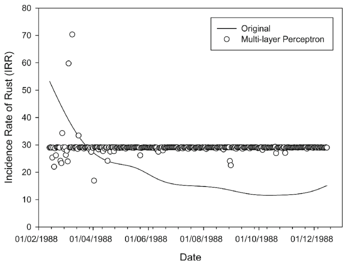
Fig. 2. Daily meteorology



M5 Regression Tree



Support Vector Regression



Multi-Layer Perceptron

Fig. 3. Meteorological variation

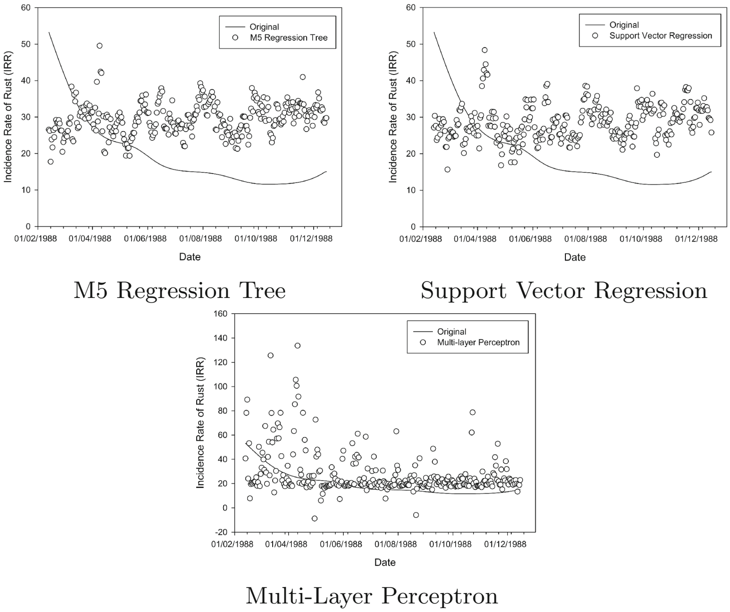


Fig. 4. Previous meteorology

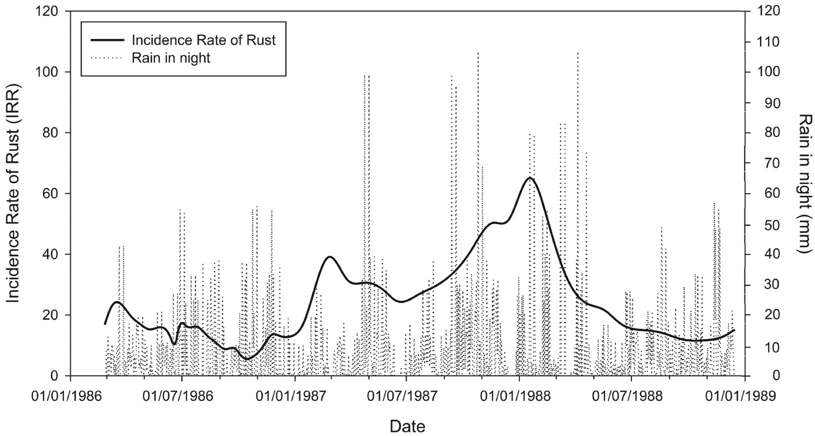


Fig. 5. IRR and Rain night for Jazmín Village

that IRR tend to peak after higher rainy seasons and then decline after performing rust control, as we can see in Fig. 5. To prove the control rust affirmation, we need data about application of fungicides on coffee plots.

In Fig. 5 we can observe higher accumulated rain in the night during seasons 17/06/1986–07/12/1986 and 20/04/1987–2/11/1987, before the two highest

Table 1. PCC and Mean Absolute Error measures

Regression models	Daily meteorology		Meteorological variation		Previous meteorology	
	r	MAE	r	MAE	r	MAE
M5	0.31	11.43%	0.12	12.04%	0.47	10.53%
SVR	0.25	11.38%	0.12	11.93%	0.48	10.02%
MLP	0.28	12.36%	0.14	12.20%	0.81	7.41%

peaks of IRR in 01/03/1987 and 18/01/1988 as stated in [23]: constant precipitations in the night is a relevant factor for germination of rust (Subsect. 2.1).

5 Conclusions and Future Works

This paper describes three data set for estimating the Incidence Rate of Rust based on meteorological time series: (i) *Daily meteorology*, (ii) *Meteorological variation*, and (iii) *Previous meteorology*. Cubic spline algorithm was used to interpolate IRR measurements during 26/02/1986 and 15/12/1988. The results show that Multi-Layer Perceptron of *Previous meteorology* is the best approximation with PCC = 0.81 and MAE = 7.41%. This approach is helpful as first approximation to test in Central American Countries where there was a crisis of coffee rust during the years 2008–2013 [2].

The hypothesis of Experimental Results presented above is that IRR tend to peak for rainy seasons and decline when it is performed rust control, nevertheless for prove it, we need to know the reliability of IRR measurement, obtaining more information about how was the Incidence Rate of Rust collected: plots, farms. Therefore it is necessary to take into account:

- *Distance between meteorological station and coffee trees:* if a weather station is away from a coffee plot, the weather measurements are inaccurate, because coffee plot can have micro-climate influenced by: coffee plot orography and properties of crop sowing such as: plant spacing, shade on coffee trees, etc. Unfortunately the weather stations are very expensive to have one per coffee plot.
- *Information about application of fungicides on coffee plots:* if fungicides are applied on coffee plots before germination of rust, the weather conditions can not be relevant factors to increase the rust incidence. We consider necessary this information to build a correct regression model based on meteorological variables.
- *Consider a margin of error in IRR measurements:* the insufficient data due to the expensive collection process that requires large expenditures of money and time [7]. The farmers must select 3 branches for each 60 coffee trees (minimum) per plot [23]. Usually one plot have 10000, 5000, or 2500 coffee trees [1], given that the maximum number of IRR measures that we can obtain for one plot are: 0.6%, 1.2%, or 2.4% respectively. Besides a coffee farm has over one coffee plot.

We propose as future work, the use Multi-Layer Perceptron of *Previous meteorology* to estimate the Incidence Rate of Rust, addressing the considerations mentioned below:

– Obtain meteorological information nearby coffee plots through forecast weather models such as meteoblue[4]. This weather models are based on NMM (Nonhydrostatic Meso-Scale Modelling) for large areas covering parts of or the entire continents, for which a complete forecast is calculated. This kind of models reduce the investment cost compared with weather stations per coffee plot.
– The properties of coffee plots are relevant information. We propose the development or the use of mobile applications for information recording of the coffee farms management as: application of fungicides, fertilizations, properties of crop sowing, etc. The main task is to encourage the farmers in the use of this kind of applications.
– Mobile technologies nowadays have a vast potential in multiple domains. To increase the measurements of IRR, we propose use mobile applications to compute the Incidence Rate of Rust based on computer vision approaches. The mobile application would analyze the pictures taken from the coffee plots, search the yellow spores on coffee trees leaves.

Acknowledgments. We thank Centro Nacional de Investigaciones de Café (Cenicafé) and Mr. Alvaro Gaitan Bustamante, PhD, for his knowledge. We also thank the Telematics Engineering Group (GIT) of the University of Cauca, and the Control Learning and Systems Optimization Group (CAOS) of the Carlos III University of Madrid, for the technical support. Finally, this work has been partially supported by Agro-Cloud project of the RICCLISA Program, and the Spanish Government (under projects TRA2011-29454-C03-03 and TRA2015-63708-R), and Colciencias for PhD scholarship granted to MsC. David Camilo Corrales.

References

1. Arcila, J., Farfan, F., Moreno, A., Salazar, L., Hincapie, E.: Sistemas de produccion de cafe en colombia. Cientific bot036, Cenicafe (2007)
2. Avelino, J., Cristancho, M., Georgiou, S., Imbach, P., Aguilar, L., Bornemann, G., Läderach, P., Anzueto, F., Hruska, A.J., Morales, C.: The coffee rust crises in colombia and central america (2008–2013): impacts, plausible causes and proposed solutions. Food Secur. **7**(2), 303–321 (2015)
3. Breiman, L., Friedman, J., Olshen, R., Stone, C.: Classification and Regression Trees. Wadsworth and Brooks, Monterey (1984)
4. Carrizosa, E., Olivares-Nadal, A.V., Ramírez-Cobo, P.: Time series interpolation via global optimization of moments fitting. Eur. J. Oper. Res. **230**(1), 97–112 (2013)
5. Cintra, M.E., Meira, C.A.A., Monard, M.C., Camargo, H.A., Rodrigues, L.H.A.: The use of fuzzy decision trees for coffee rust warning in Brazilian crops. In: 2011 11th International Conference on Intelligent Systems Design and Applications (ISDA), pp. 1347–1352, November 2011

[4] https://www.meteoblue.com.

6. Corrales, D.C., Corrales, J.C., Figueroa-Casas, A.: Towards detecting crop diseases and pest by supervised learning. Ingeniería y Universidad **19**, 207–228 (2015)
7. Corrales, D.C., Figueroa, A., Ledezma, A., Corrales, J.C.: An empirical multi-classifier for coffee rust detection in Colombian crops. In: Gervasi, O., Murgante, B., Misra, S., Gavrilova, M.L., Rocha, A.M.A.C., Torre, C., Taniar, D., Apduhan, B.O. (eds.) ICCSA 2015. LNCS, vol. 9155, pp. 60–74. Springer, Cham (2015). doi:10.1007/978-3-319-21404-7_5
8. Corrales, D.C., Casas, A.F., Ledezma, A., Corrales, J.C.: Two-level classifier ensembles for coffee rust estimation in Colombian crops. Int. J. Agric. Environ. Inf. Syst. (IJAEIS) **7**, 41–59 (2016)
9. Corrales, D.C., Ledezma, A., Peña, A.J., Hoyos, J., Figueroa, A., Corrales, J.C.: A new dataset for coffee rust detection in Colombian crops base on classifiers. Sistemas y Telemática **12**(29), 9–23 (2014)
10. Corrales, D.C., Peña, A.J., Leõn, C., Figueroa, A., Corrales, J.C.: Early warning system for coffee rust disease based on error correcting output codes: a proposal. Revista IngenierÃas Universidad de MedellÃn **13**, 57–64 (2014)
11. Friedman, J.H.: Multivariate adaptive regression splines. Ann. Stat. **19**(1), 1–67 (1991)
12. Gunn, S.R.: Support vector machines for classification and regression. Technical report (1998)
13. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The weka data mining software: an update. SIGKDD Explor. Newsl. **11**(1), 10–18 (2009)
14. Haykin, S.: Neural Networks: A Comprehensive Foundation, 2nd edn. Prentice Hall, Upper Saddle River (1998)
15. Lasso, E., Thamada, T.T., Meira, C.A.A., Corrales, J.C.: Graph patterns as representation of rules extracted from decision trees for coffee rust detection. In: Garoufallou, E., Hartley, R.J., Gaitanou, P. (eds.) MTSR 2015. CCIS, vol. 544, pp. 405–414. Springer, Cham (2015). doi:10.1007/978-3-319-24129-6_35
16. Luaces, O., Rodrigues, L.H.A., Meira, C.A.A., Bahamonde, A.: Using nondeterministic learners to alert on coffee rust disease. Expert Syst. Appl. **38**(11), 14276–14283 (2011)
17. Luaces, O., Rodrigues, L.H.A., Alves Meira, C.A., Quevedo, J.R., Bahamonde, A.: Viability of an alarm predictor for coffee rust disease using interval regression. In: García-Pedrajas, N., Herrera, F., Fyfe, C., Benítez, J.M., Ali, M. (eds.) IEA/AIE 2010. LNCS, vol. 6097, pp. 337–346. Springer, Heidelberg (2010). doi:10.1007/978-3-642-13025-0_36
18. Mohanty, P.K., Reza, M., Kumar, P., Kumar, P.: Implementation of cubic spline interpolation on parallel skeleton using pipeline model on CPU-GPU cluster. In: 2016 IEEE 6th International Conference on Advanced Computing (IACC), pp. 747–751, February 2016
19. Nutman, F.J., Roberts, F.M., Clarke, R.T.: Studies on the biology of *Hemileia vastatrix* Berk and Br. Trans. Br. Mycol. Soc. **46**(1), 27–44 (1963)
20. Perez-Ariza, C., Nicholson, A., Flores, M.: Prediction of coffee rust disease using Bayesian networks, pp. 259–266. DECSAI University of Granada (2012)
21. Quinlan, R.J.: Learning with continuous classes. In: 5th Australian Joint Conference on Artificial Intelligence, pp. 343–348. World Scientific, Singapore (1992)
22. R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria
23. Rivillas, C., Serna, C., Cristancho, M., Gaitan, A.: La Roya del Cafeto en Colombia. Impacto, manejo y costos de control. Cientific bot036, Cenicafe (2011)

24. Smola, A.J., Schölkopf, B.: A tutorial on support vector regression. Stat. Comput. **14**(3), 199–222 (2004)
25. Vapnik, V., Golowich, S.E., Smola, A.J.: Support vector method for function approximation, regression estimation and signal processing. In: Jordan, M.I., Petsche, T. (eds.) Advances in Neural Information Processing Systems, vol. 9, pp. 281–287. MIT Press, Cambridge (1997)
26. Vapnik, V.N.: Statistical Learning Theory. Wiley-Interscience, New York (1998)
27. Zivot, E., Wang, J.: Modeling Financial Time Series with S-PLUS®. IFIP. Springer, New York (2007). doi:10.1007/978-0-387-32348-0