

# Forecasting the black Sigatoka development rate: A comparison of machine learning techniques

Luis-Alexander Calvo-Valverde<sup>a,1</sup>, Mauricio Guzmán-Quesada<sup>b</sup>, José-Antonio Guzmán-Alvarez<sup>b</sup>, Pablo Alvarado-Moya<sup>c</sup>

<sup>a</sup>*DOCINADE, Instituto Tecnológico de Costa Rica, Computer Research Center, Multidisciplinar program eScience, CNCA/CeNAT, Cartago, Costa Rica*

<sup>b</sup>*Dirección de Investigaciones, Corporación Bananera Nacional S.A., Guápiles, Costa Rica*

<sup>c</sup>*DOCINADE, Instituto Tecnológico de Costa Rica, Cartago, Costa Rica*

---

## Abstract

Pending.

**Keywords:** Machine learning, Black Sigatoka, Support vector regression, Banana disease prediction, Biological warning system

---

## 1. Introduction

The black Sigatoka disease caused by the fungus *Mycosphaerella fijiensis* La bibliografía *Morelet* is the major pathological problem of banana and plantain crops in debe ser Central America, Panama, Colombia and Ecuador, as in many parts of Africa autor, anno and Asia [5].

This disease attacks the plant leaves producing a rapid deterioration of the leaf area, affects the growth and productivity of the plants due to the impairment of their photosynthetic ability causes a reduction in the quality of the fruit, and promotes premature maturation of bunches, which is the major cause of product losses associated with the black Sigatoka. Figure.1 shows three stages of this disease.

Phytopathological studies point out that precipitation, temperature, relative humidity and wind are the main climatic variables that affect its development

---

*Email address:* lualcava.sa@gmail.com (Luis-Alexander Calvo-Valverde)

<sup>1</sup>Corresponding author. (506)70104420

14 [5].

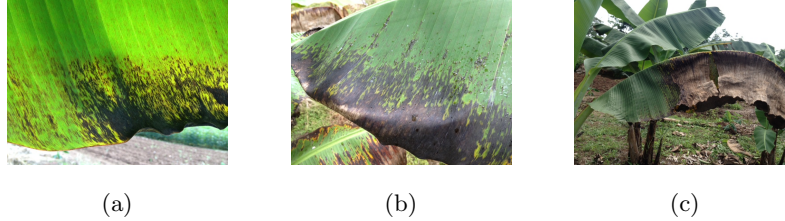


Figure 1: Examples of three disease stages of the black Sigatoka. (a) Initial stage. (b) Intermediate stage, and (c) Advanced stage.

15 In Costa Rica for the control of black Sigatoka is necessary the application  
16 of chemical fungicides. Depending of the zone of production and the weather  
17 conditions 45 - 55 cycles/year of fungicides is needed to keep this disease under  
18 control and to produce fruit with quality for exportation. This represent a cost  
19 per hectare per year ranged between \$1600 USD and \$2000 USD, about 0,64 -  
20 0.80 cents of the cost of production of a box of 18.14 kilograms. Overall, this  
21 represents 10% to 12% of the total production cost.

22 The past and present rates of disease development can in principle be used  
23 to predict its future behavior and to determine whether particular fungicide  
24 spray schedules will be able to effectively and economically control the disease  
25 Chuang and Jeger [2].

26 There are efforts to apply machine learning methods for decision-making in  
27 agriculture, including the control of crop diseases. For example, [Camargo et al.,2012]  
28 present an intelligent system for the assessment of crop disorders, [3] introduce  
29 a plant virus identification method based on neural networks with an evolu-  
30 tionary preprocessing stage, [4] summarize in their survey crop pests prediction  
31 methods using regression and machine learning approaches, while [7] present an  
32 intelligent agricultural forecasting system based on wireless sensor networks.

33 In this work, we compare five machine learning techniques (support vec-  
34 tor regression (SVR), echo state networks (ESN), ridge regression, elastic-net  
35 regression and ordinary least squares linear regression) to predict the develop-  
36 ment rate of the black Sigatoka disease.

37 The main contribution of this work is a comparison between machine learning  
38 methods to forecast black Sigatoka development rate.

FALTA COM-  
PLETAR esta  
parte

## 39 **2. Materials and methods**

### 40 *2.1. Concepts*

#### 41 *2.1.1. Black Sigatoka disease*

42 Black Sigatoka, disease caused by the fungus *Mycosphaerella fijiensis* Morelet,  
43 is the main problem phytopathologic of banana and plantain crops in Central  
44 America [5].

45 This disease attacks the leaves of plants producing a rapid deterioration of  
46 the leaf area. It affects the growth and productivity of plants by decreasing  
47 photosynthetic capacity. Also causes a reduction in quality of the fruit [5].

48 The climate has a major effect on the behavior of the black Sigatoka. Precip-  
49 itation, temperature, relative humidity and wind are the main climatic variables  
50 affecting the development of this disease [5].

#### 51 *2.1.2. Biological warning system*

52 The early warning system for black Sigatoka is an adaptation of the yellow  
53 Sigatoka warning system developed by Ganry and Meyer and modified by Ganry  
54 and Laville to use for controlling yellow Sigatoka in Cameroon. Ternesien and  
55 Fouré later improved Ganry and Laville's system. The latter system is based on  
56 weekly observations of disease symptoms on young leaves of the plant, according  
57 to Fouré's symptom (stages) descriptions. Arbitrary coefficients, based on in-  
58 cidence and severity of disease development, are used to calculate two variables:  
59 gross sum and state of evolution. Gross sum is based on the stage present and  
60 an arbitrary coefficient, which increases with the advance of the symptoms and  
61 the juvenility of the leaf. The state of evolution is calculated using the gross  
62 sum and the foliar emission period. Although threshold levels were initially  
63 suggested as a guide to spray timing, the fluctuation of these two variables was  
64 found to better define appropriate times to spray [6].

65 2.1.3. Support Vector Regression (SVR)

From the perspective of Support Vector Regression (SVR) the regression function  $y = f(s)$  for a given dataset  $D = \{(s_i, y_i)\}_{i=1}^n$ , is represented as a linear function of the form [8]:

$$f(s) = w^T s + b$$

66 where  $w$  and  $b$  are respectively the weight vector and the intercept of the model,  
67 and they are selected to find an optimal fit to the data available in  $D$ .

68 For nonlinear cases, one proceeds by mapping the input  $p$ -dimensional vec-  
69 tors via a nonlinear function  $\phi : R^p \rightarrow F$ , onto the feature space  $F$ . After  
70 nonlinear mapping, the regression function evolves to a pervasive form:

$$f(s) = w^T \phi(s) + b$$

71 SVR uses the  $\epsilon - insensitive$  loss function:

$$l = |y - f(s)|_{\epsilon} = \begin{cases} 0 & |y - f(s)| \leq \epsilon \\ |y - f(s)| - \epsilon & else \end{cases}$$

72 which ignores the error if the difference between the prediction value and  
73 the actual value is smaller than  $\epsilon$ . The  $\epsilon - insensitive$  loss function allows to  
74 find the coefficients  $w$  and  $b$  by solving a convex optimization problem, which  
75 balances the empirical error and the generalization ability. In SVR, the empirical  
76 error is measured by the loss function -insensitive and the generalization ability  
77 is measured by the Euclidean norm of  $w$  [9]. Then, the optimization problem  
78 to identify the regression model can be formulated by [8]:

$$\begin{aligned} \text{minimize} \quad & J(w, \xi_i, \xi_i^*) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n (\xi_i, \xi_i^*) \\ \text{subject to} \quad & y_i - w^T \phi(s) - b \leq \epsilon + \xi_i \\ & w^T \phi(s) + b - y_i \leq \epsilon + \xi_i^* \quad i = 1, 2, \dots, n \\ & \xi_i, \xi_i^* \geq 0 \end{aligned}$$

Falta  
completar la  
cita XXXXX  
con el libro de  
(1) SVR

79 where  $C$  denotes the penalty parameter between empirical and generalization  
 80 errors, and  $\xi_i, \xi_i^*$  are slack variables. Figure.2 shows this situation.

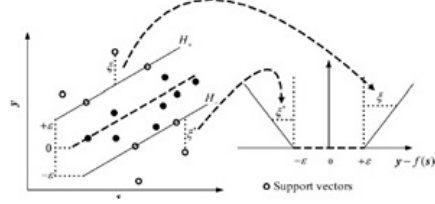


Figure 2:  $\epsilon$  - insensitive loss function[8].

The solution of this optimization problem by the Lagrange method is given by:

$$f(s) = w^T \phi(s) + b = \sum_{i=1}^n (\alpha_i - \alpha_i^*) K(s, s_i) + b$$

where  $\alpha_i - \alpha_i^*$  are the Lagrange multipliers of the optimization problems dual form and  $K(s_i, s_j)$  is the kernel function satisfying the Mercer condition, and holds:

$$K(s_i, s_j) = \langle \phi(s_i), \phi(s_j) \rangle$$

81 Operations in the kernel function  $K(s, s_i)$  are performed in the input space  
 82 rather than in the potentially high dimensional feature space of  $\phi$  [10].

#### 83 2.1.4. Ordinary least squares regression

84 This method fits a linear model with coefficients  $w = (w_1, \dots, w_p)$  to minimize  
 85 the residual sum of squares between the observed responses in the dataset, and  
 86 the responses predicted by the linear approximation. Mathematically it solves  
 87 a problem of the form [? ]:

$$\min_w \left\| Xw - y \right\|_2^2$$

88 where  $X$  denotes the features matrix.

89 According Pedregosa et al. [11] the coefficient estimates for Ordinary Least  
 90 Squares rely on the independence of the model terms. When terms are correlated

and the columns of the design matrix  $X$  have an approximate linear dependence, the design matrix becomes close to singular and as a result, the least-squares estimate becomes highly sensitive to random errors in the observed response, producing a large variance. This situation of multicollinearity can arise, for example, when data are collected without an experimental design

#### 2.1.5. Ridge regression

The ridge regression addresses some of the problems of ordinary least squares regression by imposing a penalty on the size of the coefficients. The ridge coefficients minimize a penalized residual sum of squares [11]:

$$\min_w \left\| Xw - y \right\|_2^2 + \alpha \left\| w \right\|_2^2$$

Here,  $\alpha > 0$  is a complexity parameter that controls the amount of shrinkage: the larger the value of  $\alpha$ , the greater the amount of shrinkage and thus the coefficients become more robust to collinearity.

#### 2.1.6. Elastic-Net regression

Elastic-Net is a linear regression model trained with  $L1$  and  $L2$  prior as regularizer. This combination allows for learning a sparse model where few of the weights are non-zero like Lasso, while still maintaining the regularization properties of Ridge [11]. The convex combination of  $L1$  and  $L2$  is controlled by using the *l1\_ratio* parameter.

Elastic-Net is useful when there are multiple features which are correlated with one another. Lasso is likely to pick one of these at random, while elastic-net is likely to pick both. A practical advantage of trading-off between Lasso and Ridge is it allows Elastic-Net to inherit some of Ridges stability under rotation. The objective function to minimize is [11]:

$$\min_w \frac{1}{2n_{samples}} \left\| Xw - y \right\|_2^2 + \alpha \rho \left\| w \right\|_1 + \frac{\alpha(1-\rho)}{2} \left\| w \right\|_2^2$$

### 114 2.1.7. Echo State Networks (ESN)

115 Recurrent Neural Networks (RNN) are useful for temporal patterns, but  
 116 when they are trained with backpropagation methods, they are very slow. Echo  
 117 State Network (ESN) is an alternative training method to solve that problem.  
 118 ESN is based on the observation that if a random RNN possesses certain al-  
 119 gebraic properties, training only a linear readout from it is often sufficient to  
 120 achieve excellent performance in practical applications [12]. For a given train-  
 121 ing input signal  $u(n) \in R^{N_u}$  a desired target output signal  $y^{target}(n) \in R^{N_y}$  is  
 122 known. Here  $n = 1, \dots, T$  is the discrete time and  $T$  is the number of data points  
 123 in the training dataset. The task is to learn a model with output  $y(n) \in R^{N_y}$ ,  
 124 where  $y(n)$  matches  $y^{target}(n)$  as well as possible, minimizing an error measure  
 125  $E(y, y^{target})$ , and, more importantly, generalizes well to unseen data. The un-  
 126 trained RNN part of an ESN is called a dynamical reservoir, and the resulting  
 127 states  $x(n)$  are termed echoes of its input history [13]. Finally, these signals are  
 sent to an output layer as shown in the Figure.3.

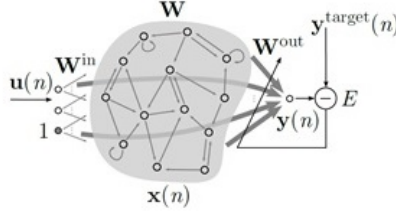


Figure 3: An echo state network [13].

128

129 The connections between the different elements of an Echo State Network  
 130 have weights randomly generated. The weights of the internal connections of  
 131 the reservoir ( $W$ ) as well as the weights of the input layer ( $W_{in}$ ), after being  
 132 generated are set statically during all stages of implementation of the algorithm.  
 133 The weights between the reservoir and the output layer ( $W_{out}$ ) are subject  
 134 to changes of a supervised learning algorithm to correct the degree of error  
 135 generated by the entire system [13].

### 136 2.1.8. Related works

137 Huang et al. [3] surveyed the development of soft computing techniques in  
138 agricultural and biological engineering, including fuzzy logic, artificial neural  
139 networks, genetic algorithms, bayesian inference and decision trees.

140 A related work, proposed by Romero [14] relies on regression models using  
141 a stepwise procedure to predict incubation and latency times of black Sigatoka.  
142 The author performed experiments on two farms located in Costa Rica (La Rita  
143 and Waldeck, the same as those used in this study but with different names).  
144 The study used data from: December 1993 to August 1995. Romero concluded  
145 that the model to predict the incubation period accounted a  $R^2$  of 69% in his  
146 observed data but it was not a good predictor when it was validated against an  
147 independent dataset (cross validation). For latency, he developed two models  
148 that accounted a  $R^2$  of 78% PONER EL VALOR OBTENIDO in the observed  
149 data, however, when validated against an independent dataset (cross validation),  
150 the model was incorrect PONER EL VALOR OBTENIDO for Weldeck, and for  
151 Rita obtained an adjusted  $R^2$  of 82%.

152 Glezakos et al. [15] proposed to use Genetic Algorithms (GA) and Neu-  
153 ral Networks (NN) to identify plant virus (Tobacco Rattle Virus (TRV) and  
154 the Cucumber Green Mottle Mosaic Virus (CGMMV)). This is achieved by  
155 the development of analytical tools of evolutionary adaptive width, propelled  
156 by Genetic Algorithms (GAs) and Neural Networks (NNs). The method was  
157 tested against some of the most commonly used classifiers in machine learning  
158 (Bayes, Trees and k-NN) via cross-validation and proved its potential towards  
159 the identification.

160 In the agricultural context, Alves et al. [16] used geoinformation techniques  
161 to develop predictive models to study the areas of risk to soybean rust in soy-  
162 bean, coffee leaf rust in coffee, and black Sigatoka in banana, considering Brazils  
163 climatic characterization and the distribution of soybean, coffee and banana  
164 crops. Temperature and rainfall data were obtained for the period from 1950  
165 to 2000, and of simulations for 2020, 2050 and 2080 using the SRES A2 cli-



mate change scenarios. Using principal components analysis, a single variable was generated based on 57 variables, in order to determine an index explaining 87%, 88% and 90% of the variability of soybean, coffee and banana crops, respectively, in municipal districts across Brazil. The climatic model was used to generate the zoning of the three plant diseases, using temperature and leaf wetness as input. Areas of favorability for the diseases were plotted against the main coffee, soybean and banana growing areas in Brazil. This methodology enabled the visualization of the changes in areas favorable for epidemics under possible future scenarios of climate change.

Other applications of machine learning methods in precision agriculture include the use of support vector regression to predict carcass weight in beef cattle in advance to the slaughter [10], machine learning assessments of soil drying for agricultural planning [17], and early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance [18].

Furthermore, there have been attempts to generate software tools. Camargo et al. [Camargo et al.,2012] presented an information system for the assessment of plant disorders (Isacrodi). They proposed that experts will attain a much better accuracy than the Isacrodi classifier, particularly when provided with samples from the affected crop. However, those cases where such expertise is not available, they suggest that Isacrodi can provide valuable support to farmers. Isacordi includes 15 crop disorders, but the black Sigatoka no is one of them. The prediction process is based on multi-class support vector machines.

Regarding the prediction of the development of the black Sigatoka with machine learning methods, Bendini et al. [19] presented a study about the risk analysis of black Sigatoka occurrence based on polynomial models. A case study was developed in a commercial banana plantation located in Jacupiranga, Brazil. It was monitored weekly during the period from February to December 2005. Data included the weekly monitoring of the diseases evolution stage, time series of meteorological data and remote sensing data. They obtained a model to estimate the evolution of the disease from satellite imagery. This model relates gray levels (NC) of the band 2 images of the Landsat-5 satellite, with the progress

197 status or disease severity (EE). The authors claim to reach an  $R^2$  of 90%.

198 Also there are works related to banana fruit. Soares et al. [20] apply two  
199 techniques: artificial neural networks (ANNs) and multiple linear regression  
200 (MLR) in banana plant to predict the yield, their results show that the neural  
201 network proved to be more accurate in forecasting the weight of the bunch in  
202 comparison to the multiple linear regressions in terms of the mean prediction-  
203 error ( $MPE = 1.40$ ), mean square deviation ( $MSD = 2.29$ ) and coefficient of  
204 determination ( $R^2 = 91\%$ ).

205 In general, the machine learning methods applied to predict the evolution  
206 of plant diseases, can be classified in two main approaches: 1) Those whose  
207 main inputs are images, and 2) Those whose main inputs are environmental  
208 and biological variables. Our study focuses in the second case.

#### 209 2.1.9. Data

210 In this work we use data acquired in two research farms of Corbana in Costa  
211 Rica: 1) 28 Millas (previously called Waldeck and located at Matina) and La  
212 Rita (located at Pococ), both in the province of Limn, Costa Rica. The banana  
213 type is Musa AAA, subgroup Cavendish, cv. Grande Naine. The Table.1 shows  
214 the variables available.

215 The value to be predicted in all cases was  $E_s$ , that is the total measure of  
216 the biological warning system.

Mas detalle

217 The data on the biological warning system are collected once a week. Al-  
218 though Corbana has meteorological stations that take data every five minutes,  
219 for these experiments, weekly averages generated by nearby stations to each of  
220 the farms were used.

aquí? o  
ampliar en los  
conceptos?

221 The time intervals used for this study were: La Rita, week 48 of 2002 to  
222 week 17 of the 2015 (647 weeks) and for 28 Miles, week 37 of 2003 to week 18  
223 of 2015 (605 weeks).

Table 1: Variables used in the study

Variable	Meaning
$T_{max}$	Max air temperature
$T_{min}$	Min air temperature
$\bar{T}_a$	Mean air temperature
$H$	Humidity
$H_{min}$	Min humidity
$H_{max}$	Max humidity
$R$	Solar radiation
$\bar{P}$	Mean precipitation
$W_{max}$	Max speed wind
$\bar{W}$	Mean speed wind
$L_2$	Biological warning system Leaf 2
$L_3$	Biological warning system Leaf 3
$L_4$	Biological warning system Leaf 4
$E_s$	Biological warning system Evolution Stage

#### 2.1.10. Data preprocessing

In 28 Miles farm, 1% of the data were missing, while in La Rita was 2.25%. To fill-in the missing values we use spline interpolation. The data collected did not exhibit outliers.

Due the fact that the variables measure meteorological or biological process, they are discretized in order to reflect trends in the data, i.e. the continuous values are not directly used. The coefficient of variation  $C_v(x)$  of each variable  $x$  was used to determine the number  $n$  of discretization levels.

$$n = \lfloor 100 C_v(x) \rfloor$$

where  $\lfloor \cdot \rfloor$  is the round operator.

Each discretization range was uniformly partitioned. Besides enabling the capture of tendencies, the discretization removes the effect of small variations

231 in the data collection, either by inaccuracies of the instruments (meteorological  
 232 variables) or by subjective bias introduced by the human who collects the data  
 233 (biological warning system). ESTO DEBE ESTAR DESCRITO EN ALGUNA  
 234 PARTE

235 Each feature was scaled to fit in a range between 0 and 1. The variable to  
 236 be predicted was not scaled.

## 237 *2.2. Evaluation criteria*

238 Although there are many types of indicators to assess the quality of the  
 239 prediction, we selected the determination coefficient ( $R^2$ ).

240 Given  $n$  records, let be  $y$  the actual value of the series,  $\hat{y}$  the predicted value  
 241 and  $\bar{y}$  the mean of the observed data.

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

$$S_e^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}$$

$$S_R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{n}$$

$$S_y^2 = S_R^2 + S_e^2$$

$$R^2 = \frac{S_R^2}{S_y^2}$$

242 This decision is supported by the widespread use in machine learning and  
 243 agriculture areas [20], [21], [22] and [23].

## 244 *2.3. Methodology*

245 The selection of methods and their parametrisation was performed in two  
 246 stages.

### 247 **Phase one**

248 In the phase one, we did ten-fold-cross-validation on a set of machine learning  
249 methods and different configurations:

- 250 • Patterns:  $n$  by  $m$ , where  $n$  from 1 to 8 and  $m$  from 1 to 2.
- 251 • Methods: support vector regression with the kernels functions: linear,  
252 RBF (Gaussian) and sigmoid; echo state networks; ordinary least squares  
253 linear regression, ridge regression and elastic-net regression.
- 254 • Variables included in the model:
  - 255 – All variables.
  - 256 – from the set  $\{\overline{T}_a, H, \overline{P}, \overline{W}\}$  use the subsets with one, two or four  
257 elements. These variables are according to experts the ones having  
258 most impact on the disease development [5].

## 259 **Phase two**

260 In the second phase, the best configurations obtained in phase one are used  
261 to validate with the last 52 and 102 weeks.

262 This second phase intents to expose how these methods behave on a consid-  
263 erable climate in the years 2014 and 2015.

### 264 *2.4. Programming environment*

265 We use the python programming language with the Integrated Development  
266 Environment (IDE) Spyder [24], particularly with the libraries pandas [25] and  
267 numpy [26]. For SVR, ridge and ordinary least squares regressions, we used  
268 sklearn [11] and for ESN the python-based code of Dr. *Lukoševičius* [13] on  
269 which the necessary were done for adjustments for the experiments of this work.  
270 The computer used a processor Intel(R) Core i7-4800MQ CPU @ 2.70GHz, 16.0  
271 GB RAM, running Windows 8 Pro.

Poner una cita  
que  
fundamente  
este ultimo  
parrafo

### 272 3. Results

### 273 4. Discussion and conclusions

### 274 5. References

- 275 [Camargo et al.,2012] Camargo, A., Molina, J., Cadena-Torres, J.,  
276 Jiménez, N., Kim, J. 2012. Intelligent systems for the assessment of  
277 crop disorders. Computers and Electronics in Agriculture(85), 1-7.  
278 doi:10.1016/j.compag.2012.02.017.
- 279 [2] Chuang, T., Jeger, M. 1987. Predicting the Rate of Development of Black  
280 Sigatoka ( *Mycosphaerella fijiensis* var. *difformis* ) Disease in Southern Tai-  
281 wan. Phytopathology, 77, 1542-1547.
- 282 [3] Huang, Y., Lan, Y., Thomson, S., Fang, A., Hoffmann, W., Lacey, R. 2010.  
283 Development of soft computing and applications in agricultural and biolog-  
284 ical engineering. Computers and Electronics in Agriculture,(71(2)), 107127.  
285 doi:10.1016/j.compag.
- 286 [4] Kim, Y., Yoo, S., Gu, Y., Lim, J., Han, D., Baik, S. 2014. Crop Pests Predic-  
287 tion Method Using Regression and Machine Learning Technology: Survey.  
288 IERI Procedia(6), 5256. doi:10.1016/j.ieri.2014.03.009.
- 289 [5] Marin Vargas, D., Romero Caldern, R. 1995. El combate de la Sigatoka  
290 Negra. Boletín Departamento de Investigaciones, Corbana Costa Rica.
- 291 [6] Marin, D., Romero, R., Guzman, M, Sutton, T. 2003. Black Sigatoka: An  
292 increasing threat to banana cultivation. Plant Disease, 87(3), 208-222.
- 293 [7] Zhao, L., He, L., Harry, W., Jin, X. 2013. Intelligent Agricultural Forecast-  
294 ing System Based on Wireless Sensor. Journal of Networks(8), 18171824.  
295 doi:10.4304/jnw.8.8.1817-1824.
- 296 [8] Wei, Z., Tao, T., ZhuoShu, D., Zio, E. (2013). A dynamic particle filter-  
297 support vector regression method for reliability prediction. Reliability Engi-  
298 neering & System Safety, 109116. doi:10.1016/j.ress.2013.05.021.

- 299 [9] Libro SVM XXXX.
- 300 [10] Alonso, J., Rodriguez Castan, ., Bahamonde, A. (2013). Support Vector Re-  
301 gression to predict carcass weight in beef cattle in advance of the slaughter.  
302 Computers and Electronics in Agriculture, 116-120.
- 303 [11] Pedregosa, F. and Varoquaux, G. and Gramfort, A. and Michel, V. and  
304 Thirion, B. and Grisel, O. and Blondel, M. and Prettenhofer, P. and Weiss,  
305 R. and Dubourg, V. and Vanderplas, J. and Passos, A. and Cournapeau,  
306 D. and Brucher, M. and Perrot, M. and Duchesnay, E. (2011) Scikit-learn:  
307 Machine Learning in Python. Journal of Machine Learning Research, vol.  
308 12, 2825–2830.
- 309 [12] Lukosevicius, M. and Jaeger, H. (2009). Reservoir computing approaches  
310 to recurrent neural network training. Computer Science Review(3), 127149.  
311 doi:10.1016/j.cosrev.2009.03.005.
- 312 [13] Lukosevicius, M. (2012). A Practical Guide to Applying Echo State Net-  
313 works. Neural Networks: Tricks of the Trade. 1-20
- 314 [14] Romero, R. (1995). Dynamics of fungicide resistant populations of My-  
315 cosphaerella fijiensis and Epidemiology of black Sigatoka of banana. Costa  
316 Rica: North Carolina State University.
- 317 [15] Glezakos, T., Moschopoulou, G., Tsiligiridis, T., Kintzios, S., Yialouris, C.  
318 (2010). Plant virus identification based on neural networks with evolutionary  
319 preprocessing. Computers and Electronics in Agriculture, 70, 263275.
- 320 [16] Alves, M., de Carvalho, L., Pozza, E., Sanches, L., Maia, J. (2011). Eco-  
321 logical zoning of soybean rust, coffee rust and banana black sigatoka based  
322 on Brazilian climate changes. Procedia Environmental Sciences, 6, 35-49.
- 323 [17] Coopersmith, E. J., Minsker, B. S., Wenzel, C. E., Gilmore, B.  
324 J. (2014). Machine learning assessments of soil drying for agricul-  
325 tural planning. Computers and Electronics in Agriculture, 104, 93104.  
326 <http://doi.org/10.1016/j.compag.2014.04.004>

- [18] Rumpf, T., Mahlein, a.-K., Steiner, U., Oerke, E.-C., Dehne, H.-W., Plmer, L. (2010). Early detection and classification of plant diseases with Support Vector Machines based on hyperspectral reflectance. *Computers and Electronics in Agriculture*, 74(1), 9199. <http://doi.org/10.1016/j.compag.2010.06.009>
- [19] Bendini, H., Moraes, W., da Silva, S., Tezuka, E., Cruvinel, P. (2013). Anlise de risco da ocorrência de Sigatoka-negra baseada em modelos polinomiais: um estudo de caso. *Tropical Plant Pathology*, 38, 035-043.
- [20] Soares, J., Pasqual, M., Lacerda, W., Silva, S., Donato, S. (2014). Comparison of techniques used in the prediction of yield in banana plants. *Scientia Horticulturae journal*, 167, 84-90.
- [21] Soares, J., Pasqual, M., Lacerda, W. (2013). Utilization of artificial neural networks in the prediction of the bunchesweight in banana plants. *Scientia Horticulturae*(155), 24-29.
- [22] Ibrahim, N. and Wibowo, A. (2014). Time Series Support Vector Regression with Missing Data Treatment Based Variables Selection for Water Level Prediction of Galas River in Kelantan Malaysia. *International Journal of Applied Research in Engineering and Science*, 3, 25-36.
- [23] Demir, B. and Bruzzone, L. (2014). A multiple criteria active learning method for support vector regression. *Pattern Recognition*, 25582567. doi:10.1016/j.patcog.2014.02.001
- [24] Continuum Analytics. (2015). Anaconda. Retrieved from <https://www.continuum.io/>
- [25] McKinney W. (2010). Data Structures for Statistical Computing in Python, *Proceedings of the 9th Python in Science Conference*, 51-56
- [26] van der Walt S., Colbert C. and Varoquaux G. (2011). The NumPy Array: A Structure for Efficient Numerical Computation, *Computing in Science & Engineering*, 13, 22-30 (2011), DOI:10.1109/MCSE.2011.37