

## ORIGINAL RESEARCH

# Predicting the rates of cross-pollination between GM and non-GM crops using RBFNN with SVM and bootstrap approach

Chien-Pang Lee<sup>1</sup>, Bo-Jein Kuo<sup>2</sup>, Wen-Shin Lin<sup>3</sup>, Kuo-Chung Chu<sup>4</sup>

1. Department of Information Management, Da-Yeh University, Taiwan. 2. Biometry Division, Department of Agronomy, National Chung Hsing University, Taiwan. 3. Department of Plant Industry, National Pingtung University of Science and Technology, Taiwan. 4. Department of Information Management, National Taipei University of Nursing and Health Sciences, Taiwan.

**Correspondence:** Chien-Pang Lee. Address: No. 168, University Rd., Dacun, Changhua 51591, Taiwan. Telephone: 886-485-118-883-132. E-mail: cplee@mail.dyu.edu.tw.

**Received:** March 26, 2012

**Accepted:** June 7, 2012

**Published:** September 1, 2012

**DOI:** 10.5430/air.v1n1p84

**URL:** <http://dx.doi.org/10.5430/air.v1n1p84>

## Abstract

Recently, genetically modified (GM) technology has been successfully used to reduce the cost and to enhance the profit in agriculture. Although GM technology brings many benefits for non-food crops, people still misgive the effects of GM products for the health and the environment. Furthermore, GM crops might affect food (non-GM) crops in the open environment. Hence, how to find strategies for the coexistence of GM and non-GM crops are become a popular issue. One of the strategies is to determine a befitting distance between GM and non-GM crops to reduce the cross-pollination occurred by predicting the rate of cross-pollination. Owing to most of the existing methods for predicting the cross-pollination rates of non-GM crops are only based on the distance between GM and non-GM crops. To counter this problem, we propose a hybrid method, which is composed of radial basis function neural network (RBFNN), support vector machine (SVM) and bootstrap, to apply in this issue. The proposed method includes three specificities. (a) The proposed method reduces the effect of imbalance class problem. (b) The proposed method uses more variables, which are effect the cross-pollination rates, for prediction to enhance the prediction accuracy. (c) The proposed method searches relevant samples to reduce execution time and enhance the prediction accuracy. The results show the performance of our method is better than the existing methods in terms of the root mean square error (RMSE) in prediction and the correlation coefficient between the actual and the predicted cross-pollination rates.

## Key words

Genetically modified, Radial basis function neural network, Support vector machine, Bootstrap, Imbalanced class distribution problem

## 1 Introduction

Recently, biotechnology has been successfully applied in agriculture to reduce the cost and to enhance the profit of agricultural products. Since the U.S. open the field cultivation of genetically modified (GM) crops in 1996, the acreage of GM crops have been growing. According to statistics in 2010, the main GM crops of the world are soybean (50%), maize

(31%), cotton (14%), and canola (5%)<sup>[1]</sup>. The development of GM crops bring many benefits, such as (1) crops have become more insect-resistant; (2) food crops can be added vitamin or trace metal to benefit the human health; (3) The quality of non-food crops are improved for biomass energy or papermaking<sup>[2]</sup>. Although GM technology brings many benefits for non-food crops, people still misgive the effect of GM products for the health and the environment. Furthermore, GM crops might affect food (non-GM) crops in the open environment. Consequently, to avoid food crops are affected by GM crops, the strategies for the coexistence of GM and non-GM crop have become the popular issue in the world<sup>[3]</sup>.

Due to the cross-pollination might occur between GM and conventional (or organic) crops in the open environment, the tolerance threshold of the cross-pollination rate for non-GM agricultural products is set to 0.9% in the European Union (EU). Consequently, one of the strategies for the coexistence of GM and non-GM crops is predicting the cross-pollination rates of non-GM crops at different distance between GM and non-GM crops. Subsequently, government or researcher can determine a befitting distance between GM and non-GM crops to reduce cross-pollination occurred by using a model with higher accuracy in predicting cross-pollination rates. Owing to cross-pollination rates are associated with distance between GM and non-GM crops, most of the existing methods for predicting the cross-pollination rates of crops, which are only based on the distance between GM and non-GM crops. For example, due to the distribution of cross-pollination rates is close to exponential distribution, Ma et al.<sup>[4]</sup> proposed a non-linear model, which is modified from exponential distribution, to predict cross-pollination rates in 2004. Ma et al.'s model was only based on the distance between GM and non-GM crops. Subsequently, Gustafson et al.<sup>[5]</sup> used a power function to predict cross-pollination rates in 2006. In Gustafson et al.'s model, the distance was the variable of power function and used to estimate the other parameters of Gustafson et al.'s model. In 2008, Della et al.<sup>[6]</sup> proposed another non-linear model, which is  $c10^{k\sqrt{x}}$ . In Della et al.'s model,  $x$  denoted the distance;  $c$  and  $k$  were two parameters, which were estimated by distance.

Obviously, the well-known models for predicting cross-pollination rates only use distance between GM and non-GM crops. However, according to the literatures<sup>[7, 8]</sup>, the cross-pollination occurred should be affected by many variables (factors) such as flowering time, wind speed, wind direction, sunshine-hour, and rainfall. For this reason, this paper wants to propose a method, which considers many variables (factors) to predict the rates of cross-pollination. Our proposed model is called RBFNN with SVM and bootstrap approach which is composed of radial basis function neural network (RBFNN), support vector machine (SVM) and bootstrap method. The main idea of the proposed model contains three parts. Firstly, the data of cross-pollination rates are classified into the categories of "GM crops" and "non-GM crops" according to the rate of cross-pollination. Subsequently, the training data are selected by the proportion of two categories to avoid the problem of imbalance class. Secondly, SVM and Euclidean distance to select relevant samples, which are similar to the testing sample, as training samples to enhance the prediction accuracy and reduce the execution time. Finally, the proposed model uses RBFNN to build the prediction model and uses bootstrap method to ensure the performance of predicting cross-pollination rates.

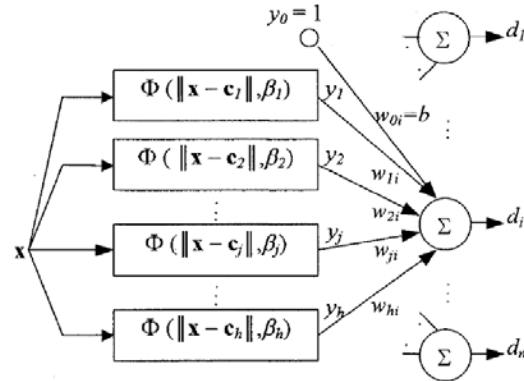
The remainder of this paper is organized as follows. Section 2 briefly reviews the definitions of RBFNN and SVM. Section 3 introduces the details of the proposed method. In Section 4, we report the experiment and results and Section 5 concludes this paper.

## 2 Related methods

### 2.1 Radial basis function neural network

The concept of artificial neural network (ANN) model was first introduced in 1950s<sup>[9]</sup>. Since many artificial neural network models have been proposed. Among them, the feed-forward neural network and the radial basis function neural network (RBFNN) are most well-known artificial neural network model<sup>[9, 10]</sup>. Furthermore, RBFNN has become a popular

and powerful neural network for predicting in recent years. In comparison to other neural network algorithms like MLP, it uses less computation time for training as a result of its simple framework and has a more compact structure<sup>[11]</sup>. For these reasons, this paper uses RBFNN to predict the rate of cross-pollination. Figure 1 shows the structure of RBFNN. Two typical choices for RBF functions are shown in the following<sup>[12]</sup>.



**Figure 1.** The structure of radial basis function neural network<sup>[13]</sup>

Thin-plate-spline function:

$$\phi(r) = r^2 \log(r) \tag{1}$$

Gaussian function:

$$\phi(r, \beta) = \exp(-r^2 / \beta^2), \tag{2}$$

where  $r$  is calculated by the Euclidean distance as shown in Eq. (3):

$$r_j = \sqrt{\sum_{i=1}^n (x_i - c_j)^2}, \tag{3}$$

where  $x_i$  is the  $i$ th input and  $c_j$  is the center of the  $j$ th RBF. In general, RBF centers are chosen randomly or from some arbitrary data points. Then, the output can be calculated by Eq. (4) and Eq. (5).

$$d_i = w_{0i} + \sum_{j=1}^h w_{ji} y_j, \tag{4}$$

$$y_j = \phi(\|x - c_j\|, \beta_j), \quad i = 1, \dots, m \text{ and } j = 1, \dots, h, \tag{5}$$

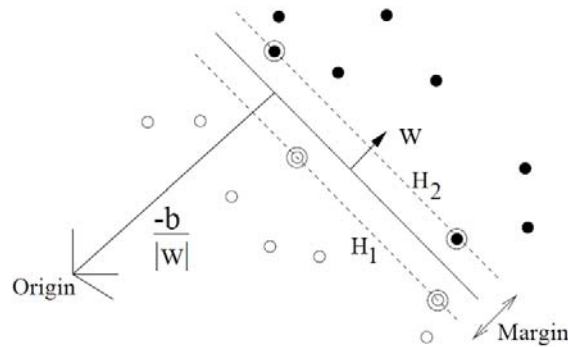
where  $d_i$  is the  $i$ th output;  $w_{0i}$  is the biasing term;  $h$  is the number of hidden neurons,  $w_{ji}$  is the weight between the  $j$ th hidden neuron and the  $i$ th output,  $\beta_j$  is the real constant called spread factor of the  $j$ th hidden neuron.

## 2.2 Support vector machine

Support vector machine (SVM) was first introduced in 1992<sup>[14]</sup>. SVM is a supervised machine learning classifier. SVM has become a popular and powerful classifier in the recent years. SVM builds a model that predicts whether a new

example falls into one category or the other. The basic binary SVM is introduced in the following. The original idea of the binary SVM is to use a linear separating hyper-plane (decision boundary) to maximize the distance between two classes<sup>[15]</sup> as shown in Fig. 2. The decision function of the binary SVM is presented as  $f(X) = W^T X + b$ , where  $W = [w_1, w_2, \dots, w_n]^T$  is the weight vector,  $b$  is a scalar  $n$  denotes the number of variables. The decision function is obtained by solving the following optimization problem.

$$\begin{aligned} & \text{minimize } P(w, \xi) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i, \\ & \text{subject to } y_i [w^T x_i + b] \geq 1 - \xi_i, \\ & \quad y_i \in \{-1, 1\}, \xi_i \geq 0, \quad i = 1, 2, \dots, n, \end{aligned} \quad (6)$$



**Figure 2.** Linear separating hyperplanes for the separable case<sup>[16]</sup>

Where parameter  $C$  is a trade-off between training accuracy and generalization. The solution of this problem is given by solving the corresponding dual problem, a quadratic programming problem with  $n$  variables<sup>[17]</sup>.

### 3 RBFNN with SVM and bootstrap approach

The main idea of the proposed model consists of three steps as described as follows. Part 1: classify dataset: the dataset is classified into the categories of “*GM crops*” and “*non-GM crops*” according to the rate of cross-pollination. Then, the training data are selected by the proportion of two categories to avoid the imbalance class problem. Part 2: searching similar training data: the proposed model uses SVM and Euclidean distance to select relevant samples, which are similar to the testing sample, for building prediction model to enhance the prediction accuracy and reduce the execution time. Part 3: predicting cross-pollination rate: the proposed model uses the selected relevant samples to build RBFNN model for predicting the rates of cross-pollination and use bootstrap method to ensure the performance of predicting cross-pollination rates. Fig. 3 shows the flowchart of the proposed model, and the details are introduced in the following.

#### Step 1: classify dataset

As described as Section 1, the main effect of the cross-pollination occurred is the distance between GM and non-GM crops. Moreover, most of samples are with smaller rates of cross-pollination. Consequently, when most of the training dataset are with smaller rates of cross-pollination, the built model could not predict a sample with higher rate of cross-pollination. This situation is called imbalance class problem. To avoid this problem, we, firstly, perform data pre-processing before predicting.

Each of data is classified into the categories of “GM crops” and “non-GM crops” according to the rate of cross-pollination. To follow the definition of EU for GM and non-GM crops, the criterion of this paper for classifying dataset is set to 0.9%. If a sample’s cross-pollination rate is greater than 0.9%, the sample will be classified into the category of “GM crops”. In contrast, the sample will be classified into the category of “non-GM crops”. Subsequently, the training data in this paper are selected according to the proportion of the result of the classified categories. For example, 1,000 and 100 samples are classified into the category of “non-GM crops” and that of “GM crops”, respectively.  $100 \times 70\% = 70$  samples belong to the category of “GM crops” and  $70 \times (1+50\%) = 105$  samples belong to the category of “non-GM crops” are selected as training dataset for the proposed model. Remainder samples are testing dataset. Accordingly, it can solve the imbalance class problem by this step.

**Step 2: the relevant samples selection**

After selecting the training data, we want to select the similar samples, which are called relevant samples, for the testing sample to enhance the prediction accuracy. Hence, each of testing sample has its own model, which is a specific model for predicting the cross-pollination rate of the test sample. Consequently, we use the following three steps to select relevant samples.

**Step 2.1: constructing the proportion table for selecting relevant samples**

Consequently, we use SVM to construct a table of classified results of the training data as shown as Table 1. Next, we construct a proportion table for selecting relevant samples, as shown as Table 2, based on Table 1.

**Table 1.** A table of classified results of the training data

		Actual		
		Non-GM crops	GM crops	total
Predicted	Non-GM crops	<i>a</i>	<i>b</i>	<i>a+b</i>
	GM crops	<i>c</i>	<i>d</i>	<i>c+d</i>
	total	<i>a+c</i>	<i>b+d</i>	

**Table 2.** A proportion table for selecting relevant samples.

		Actual	
		Non-GM crops	GM crops
Predicted	Non-GM crops	$P_{11} = a/(a+b)$	$P_{12} = 1 - P_{11}$
	GM crops	$P_{21} = c/(c+d)$	$P_{22} = 1 - P_{21}$

**Step 2.2: calculate the similarity between the testing sample and the training data**

This step uses the Euclidean distance to calculate the similarities between a testing sample and each of training data as shown in Eq. (7). In Eq. (7),  $T_j$  denotes the  $j$ th variable of the testing sample;  $X_{ij}$  denotes the  $j$ th variable of the  $i$ th training data;  $n$  denotes the number of variables. The similarities between the testing sample and each training data are sorted as shown in Table 3.

$$similarity_i = \sqrt{\sum_{j=1}^n T_j - X_{ij}}, \tag{7}$$

**Table 3.** The sorted similarities table

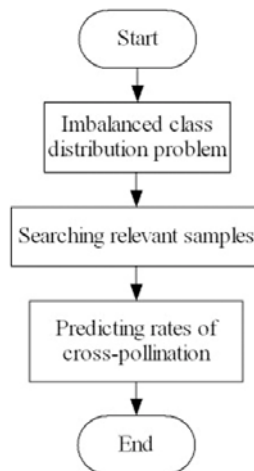
Rank	Non-GM crops	GM crops
1	$ED_{N1}$	$ED_{G1}$
2	$ED_{N2}$	$ED_{G2}$
...	...	...
$n$	$ED_{Nn}$	$ED_{Gn}$
...	...	...
$m$	$ED_{Nm}$	

**Step 2.3: searching relevant samples for prediction**

In this step, SVM is used to classify the testing sample. Then, the relevant samples are selected by the predicted category of the test sample. For example, if a testing sample is classified into the category of “non-GM crops”, the first  $200 \times P_{11}$  ( $P_{11}$  is a proportion in Table 2) samples are selected from training dataset belong to the category of “non-GM crops” in Table 3 and the first  $200 \times P_{12}$  ( $P_{12}$  is a proportion in Table 2) samples are selected from training dataset belong to the category of “GM crops” in Table 3. After selecting 200 relevant samples, they are used to build a specific model for predicting the cross-pollination rate of the test sample in Step 3.

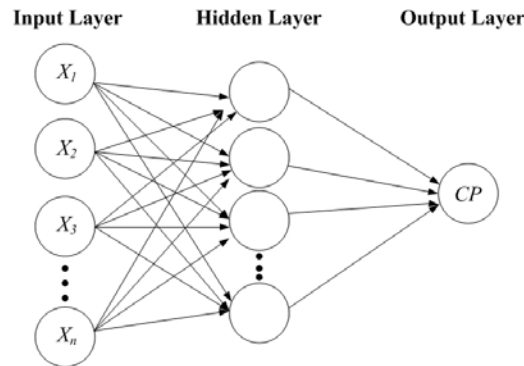
**Step 3: predict the rate of cross-pollination**

After having relevant samples, we use them to train RBFNN model for predicting cross-pollination rates. Figure 3 shows the framework of RBFNN model for predicting cross-pollination rates. In Fig. 4, the nodes of input layer are the variables ( $X_1, X_2, X_3, \dots, X_n$ ) of the data and the node of output layer is the rate of cross-pollination ( $CP$ ). Finally, we perform predicting cross-pollination rate by feeding the testing sample into the trained RBFNN model.

**Figure 3.** The flowchart of RBFNN with SVM bootstrap approach**Step 4: bootstrap approach**

In statistics, bootstrap is a resampling with replacement for small sample size to reduce the variance of the prediction error<sup>[18]</sup>. Hence, to enhance the stability of our proposed method, Step 3 is performed 100 times bootstrap method in this paper. Consequently, the final predicted cross-pollination rate is the average predicted cross-pollination rate from the results of bootstrap method. Note that an iteration of the above procedure (Step 1 through Step 4) predicts only one cross-pollination rate of a testing sample.

Note that the programming of this paper is coded by the statistical software R. Furthermore, “e1071” package of R is used to build SVM model, and “RSSNS” package of R is used to build RBFNN model.



**Figure 4.** The framework of the proposed model for prediction cross-pollination rates

## 4 Experiment and results

### 4.1 Dataset

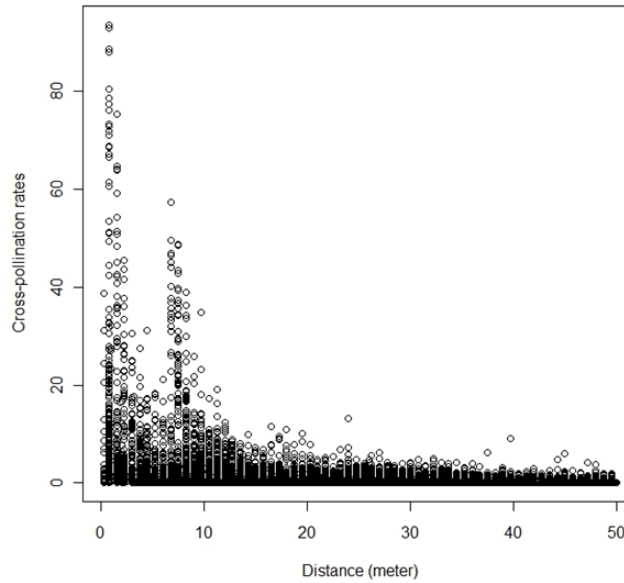
The dataset of this paper are the first and second harvests of maize, contain in 2009 to 2011, from Taichung District Agricultural Improvement Station and Potzu Branch Station in Tainan District Agricultural Improvement Station. The data contains 17,482 samples and 8 variables. Note that the descriptions of the variables are shown in Table 4. The response variable of this paper is “CP”, and the remainder variables are explanatory variables. The maximum, average, and standard deviation of the cross-pollination rates are 93.7%, 1.069%, and 4.4716 %, respectively, as shown in Table 5. The trend of the cross-pollination rates at different the distance from pollen source is shown in Fig. 5. To enhance the availability of our proposed method, we divide data into five sets according to the distance from the GM crops. The data distributions of the cross-pollination rates at different distance are shown in Table 5. Subsequently, 66% samples are selected into training dataset, and the remainder samples are testing dataset. Accordingly, the training dataset is used to train and build prediction model, and the testing dataset is used to verify the performance of our proposed method.

**Table 4.** The descriptions of the variables

Variable	Description	Variable	Description
CP	The rates of cross-pollination	MWS	The average wind speed (m/s) of month
MT	The average temperature (°C) of month	RAIN	The average rainfall (mm) of month
Distance	The distance between GM and non-GM crops	theta	The angle theta between GM and non-GM crops
MH	The average relative humidity (%) of month	SUN	The average sunshine-hour of month

**Table 5.** The data distribution of the cross-pollination rates at different distance

	Mean (%)	Standard deviation (%)	Maximum (%)	Median (%)
0-10 meter	2.9330	8.3923	93.7	0.0948
0-20 meter	1.7770	6.1243	93.7	0
0-30 meter	1.3230	5.1185	93.7	0
0-40 meter	1.1840	4.7542	93.7	0
0-50 meter	1.0690	4.4716	93.7	0



**Figure 5.** The trend of the cross-pollination rates at different distance.

## 4.2 Performance

In this section, we report our experimental results and compare to the existing methods. The existing methods, which are exponential, log/log, and log/quart, are introduced in the literatures<sup>[4-6]</sup> as represented in Eq. (8), Eq. (9), and Eq. (10), respectively.

$$cp = c \exp(kx), \quad (8)$$

$$cp = cx^k, \quad (9)$$

$$cp = c10^{k\sqrt{x}}, \quad (10)$$

where  $cp$  denotes the rate of cross-pollination;  $x$  denotes the distance between GM and non-GM crops;  $c$  and  $k$  are two parameters of equations.

Due to the root mean square error (RMSE) is a main performance measure for predicting<sup>[9-19]</sup>, this paper uses RMSE, as shown as Eq. (11), for comparing the performance between models. In Eq. (11),  $CP_i$  denotes the actual cross-pollination rate of the  $i$ th testing sample;  $\overline{CP}_i$  denotes the predicted cross-pollination rate of the  $i$ th testing sample.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (CP_i - \overline{CP}_i)^2}{n}}. \quad (11)$$

The RMSE (%) of the proposed model and the existing models at different distance are shown in Table 6 and Figure 6. For 0-10 meter, the RMSEs (%) of the existing models are similar. In contrast, the RMSE (%) of the proposed model is better



than the existing models. Obviously, Table 6 and Figure 6 show that the RMSE (%) of the proposed model is the smallest at each distance. The RMSE (%) of the proposed method is less than that of the existing methods about 30%.

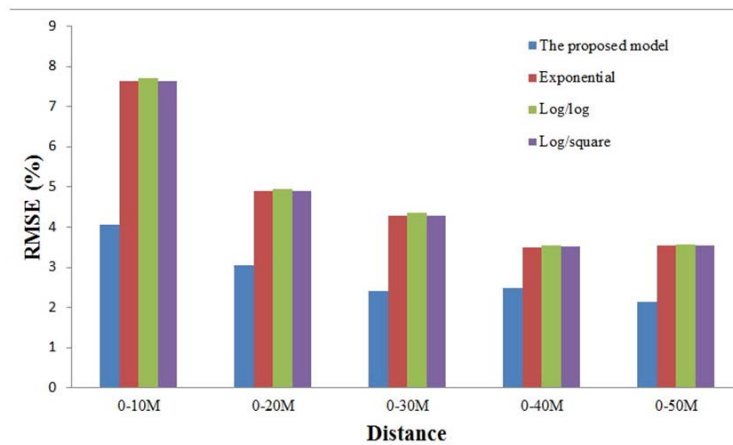


Figure 6. RMSE (%) of the proposed model and the existing methods

Table 6. RMSE (%) of the proposed model and the existing methods

	The proposed model	Exponential	Log/log	Log/square
0-10M	4.0645*	7.6300	7.6978	7.6442
0-20M	3.0627*	4.9085	4.9533	4.8982
0-30M	2.4086*	4.2800	4.3648	4.2822
0-40M	2.4798*	3.4923	3.5473	3.5082
0-50M	2.1429*	3.5488	3.5696	3.5384

\*\*\* denotes the minimum RMSE

Table 7 shows the coefficient of correlation between the actual and the predicted cross-pollination rates for the proposed model and the existing methods at different distance. In table 7, the coefficients of correlation of the proposed method are significant greater than that of the existing methods. The average of correlation coefficient of our proposed method is greater than 0.75. In contrast, the averages of correlation coefficients of the existing models are about 0.2. Figure 7 shows the scatter plot of the actual and the predicted cross-pollination rates for the distance at 0-10 meter.

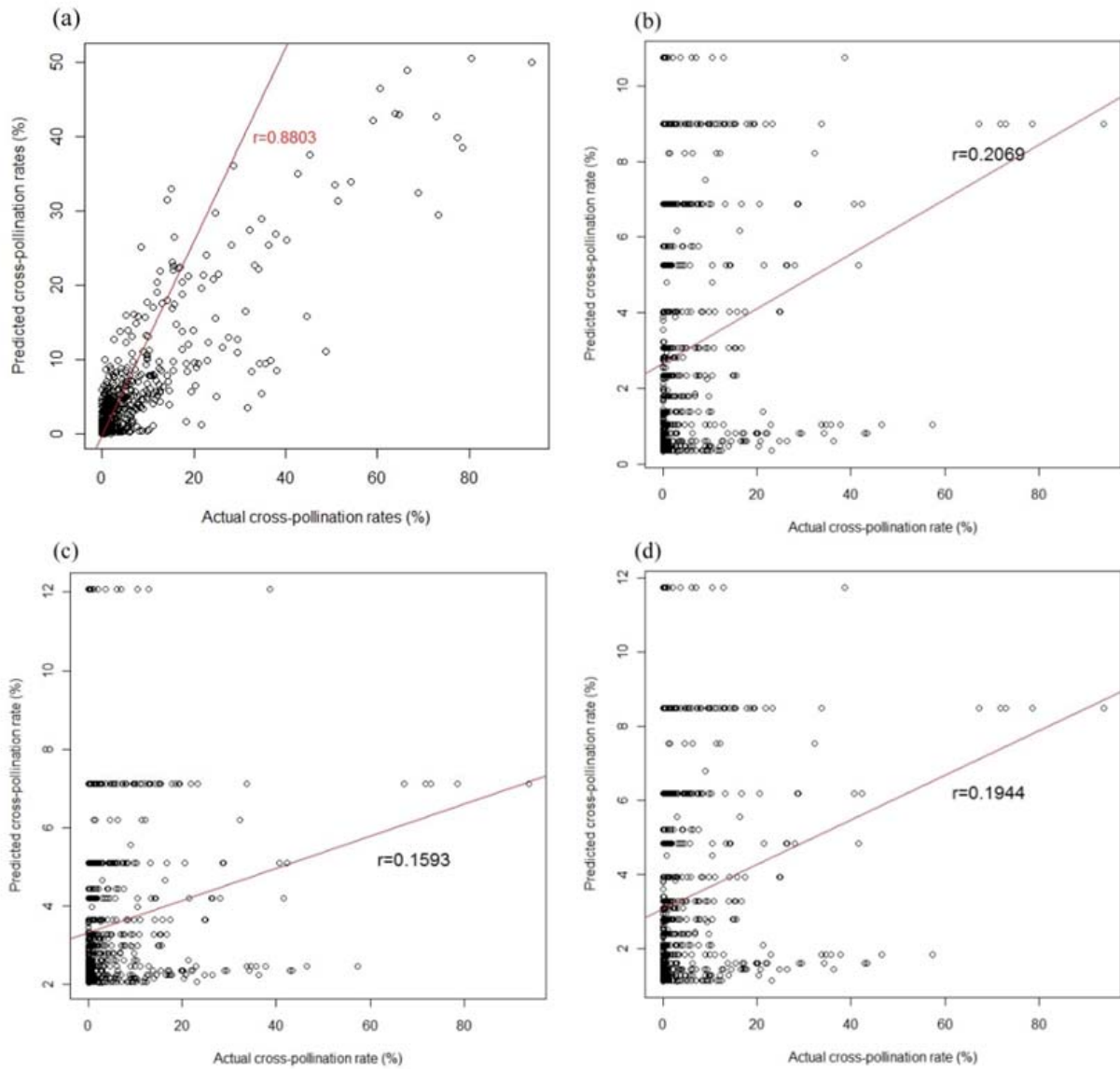
Table 7. The coefficient of correlation between the actual and the predicted cross-pollination rates

	The proposed model	Exponential	Log/log	Log/square
0-10M	0.8803*	0.2069	0.1593	0.1944
0-20M	0.8257*	0.2525	0.2370	0.2626
0-30M	0.7961*	0.2484	0.2294	0.2600
0-40M	0.7860*	0.2291	0.2005	0.2302
0-50M	0.7693*	0.2311	0.2238	0.2411

\*\*\* denotes the maximum coefficient of correlation

To verify the performance of the proposed model is significant better than the other models, we test the different of RMSEs (%) between models. Table 8 shows the results (*p*-value) of testing at 0-50 meter. In Table 8, the *p*-values of the proposed model are small than 0.001, namely, they are significant better than the existing models at significant level  $\alpha=0.05$ . The *p*-values between the other existing models are greater than 0.9, namely, they are insignificant different. In

short, the performance of the proposed model is better than the existing models in terms of RMSE (%) and the coefficient of correlation, especially the correlation coefficient of our proposed method is greater than 0.88 ( $r=0.8803$ ) for the distance at 0-10 meter. Hence, the proposed model can be applied for prediction the rates of cross-pollination.



**Figure 7.** The scatter plot of the actual vs. the predicted cross-pollination rates for the distance at 0-10 meter. (a) The proposed model; (b) Exponential; (c) Log/log; (d) Log/square

**Table 8.** The test results of RMSE at 0-50 meter

	The proposed model	Exponential	Log/log
Exponential	<0.001		
Log/log	<0.001	0.9375	
Log/square	<0.001	0.9678	0.9055

## 5 Conclusion

Recently, biotechnology has been successfully applied in agriculture. The GM technology is one of the important biotechnologies in agriculture. However, people still misgive the effect of GM crops products for the health and the environment. Hence, the strategies for coexistence of GM and non-GM crops have become the popular research issue. The rate of cross-pollination prediction is one of the strategies for the coexistence of GM and non-GM crops. This paper proposes RBFNN with SVM and bootstrap approach to predict the rates of cross-pollination. The proposed method includes three specificities, such as (a) the proposed method reduces the effect of imbalance class problem. (b) The proposed method uses more variables, which are effect the cross-pollination rates, for prediction to enhance the prediction accuracy. (c) The proposed method searches relevant samples to reduce execution time and enhance the prediction accuracy. According to the results, the performance of the proposed model is significant better than that of the existing methods in terms of RMSE (%) and the correlation coefficient. Therefore, the proposed model could be successful applied for the cross-pollination rates prediction. In the further, governments or researchers can use the proposed model to determine a fitness distance for reducing the effect between GM and non-GM crops. Consequently, governments or researchers can make a decision for the policy of the coexistence of GM and non-GM crops according to the results of the proposed model.

## References

- [1] James, C. Global status of commercialized biotech/GM Crops: 2010. Ithaca, NY: ISAAA. 2011.
- [2] Davison, J., & Bertheau, Y. EU regulations on the traceability and detection of GMOs: difficulties in interpretation, implementation and compliance. *CAB Reviews*. 2007; 2 (77): 1-14. <http://dx.doi.org/10.1079/PAVSNNR20072077>
- [3] Chan, L. F., Din, M. Y., Wu, M. T., Wu, H. C., Wei, M. L., & Lu, H. Y. Co-existence model and isolation distance in genetically modified crops. *Scientific Agriculture*. 2008; 56 (1-3): 35-47.
- [4] Ma, B. L., Subedi, K. D., & Reid, L. M. Extent of cross-fertilization in maize by pollen from neighboring transgenic hybrids. *Crop Science*. 2004; 44 (4): 1273-1282. <http://dx.doi.org/10.2135/cropsci2004.1273>
- [5] Gustafson, D. I., Brants, I. O., Horak, M. J., Remund, K. M., Rosenbaum, E. W., & Soteres, J. K. Empirical modeling of genetically modified maize grain production practices to achieve European Union labeling thresholds. *Crop Science*. 2006; 46 (5): 2133-2140. <http://dx.doi.org/10.2135/cropsci2006.01.0060>
- [6] Della Porta, G., Ederle, D., Bucchini, L., Prandi, M., Verderio, A., & Pozzi, C. Maize pollen mediated gene flow in the Po valley (Italy): Source-recipient distance and effect of flowering time. *European Journal of Agronomy*. 2008; 28 (3): 255-265. <http://dx.doi.org/10.1016/j.eja.2007.07.009>
- [7] Allnut, T. R., Dwyer, M., McMillan, J., Henry, C., & Langrell, S. Sampling and modeling for the quantification of adventitious genetically modified presence in maize. *Journal of Agricultural and Food Chemistry*. 2008; 56 (9): 3232-3237. PMID:18419127 <http://dx.doi.org/10.1021/jf800048q>
- [8] Messeguer, J., Peñas, G., Ballester, J., Bas, M., Serra, J., Salvia, J., et al. Pollen-mediated gene flow in maize in real situations of coexistence. *Plant Biotechnology Journal*. 2006; 4 (6): 633-645. PMID:17309734 <http://dx.doi.org/10.1111/j.1467-7652.2006.00207.x>
- [9] Leu, Y., Lee, C.-P., & Jou, Y. Z. A distance-based fuzzy time series model for exchange rates forecasting. *Expert Systems with Applications*. 2009; 36 (4): 8107-8114. <http://dx.doi.org/10.1016/j.eswa.2008.10.034>
- [10] Panda, C., & Narasimhan, V. Forecasting exchange rate better with artificial neural network. *Journal of Policy Modeling*. 2007; 29 (2): 227-236. <http://dx.doi.org/10.1016/j.jpolmod.2006.01.005>
- [11] Kavousifard, A., & Samet, H. Consideration effect of uncertainty in power system reliability indices using radial basis function network and fuzzy logic theory. *Neurocomputing*. 2011; 74 (17): 3420-3427. <http://dx.doi.org/10.1016/j.neucom.2011.05.017>
- [12] Chen, S., Cowan, C. F. N., & Grant, P. M. Orthogonal least squares learning algorithm for radial basis function networks. *IEEE Transactions on Neural Networks*. 1991; 2 (2): 302-309. PMID:18276384 <http://dx.doi.org/10.1109/72.80341>
- [13] Segal, R., Kothari, M. L., & Madhani, S. Radial basis function (RBF) network adaptive power system stabilizer. *IEEE Transactions on Power Systems*. 2000; 15 (2): 722-727. <http://dx.doi.org/10.1109/59.867165>
- [14] Boser, B.E., Guyon, I.M., & Vapnik V.N. A training algorithm for optimal margin classifiers. *Proceedings of the fifth annual workshop on Computational learning theory*. 1992; 144-152. <http://dx.doi.org/10.1145/130385.130401>

- [15] Huang, H., & Chang F. ESVM: Evolutionary support vector machine for automatic feature selection and classification of microarray data. *Biosystems*. 2007; 90 (2): 516-528. PMID:17280775 <http://dx.doi.org/10.1016/j.biosystems.2006.12.003>
- [16] Burges, C.J.C. A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*. 1998; 2: 121-167. <http://dx.doi.org/10.1023/A:1009715923555>
- [17] Zhou, X., & Tuck, D.P. MSVM-RFE: extensions of SVM-RFE for multiclass gene selection on DNA microarray data. *Bioinformatics*. 2007; 23 (9): 1106-1114. PMID:17494773 <http://dx.doi.org/10.1093/bioinformatics/btm036>
- [18] Buja , A., & Stuetzle, W. Observations on bagging. *Statistica Sinica*. 2006; 16: 323-351.
- [19] Leu, Y., Lee, C.-P., & Hung, C.-C. A fuzzy time series-based neural network approach to option price forecasting. *Lecture Notes in Computer Science*. 2010; 5990: 360-369. [http://dx.doi.org/10.1007/978-3-642-12145-6\\_37](http://dx.doi.org/10.1007/978-3-642-12145-6_37)