

Moisture Content Prediction of Dried Longan Aril from Dielectric Constant Using Multilayer Perceptrons and Support Vector Regression

¹Sanong Amaroek, ^{2,3}Nipon Theera-Umpon,
³Kittichai Wantanajittikul and ^{3,4}Sansanee Auephanwiriyakul
¹Postharvest Technology Institute,
²Department of Electrical Engineering, Faculty of Engineering,
³Biomedical Engineering Center,
⁴Department of Computer Engineering, Faculty of Engineering,
Chiang Mai University, Chiang Mai 50200 Thailand

Abstract: Problem statement: Estimation of moisture contents of dried food products from their dielectric constants was an important step in moisture measurement systems. The regression models that provide good prediction performance are desirable. **Approach:** The Multilayer Perceptrons (MLP) and Support Vector Regression (SVR) were applied in this research to predict the moisture contents of dried longan arils from their dielectric constants. The data set was collected from 1500 samples of dried longan aril with five different moisture contents of 10, 14, 18, 22 and 25% Wet basis (Wb.) Dielectric constant of dried longan aril was measured by using our previously proposed electrical capacitance-based system. The results from the MLP and SVR models were compared to that from the linear regression and polynomial regression models. To take into account the generalization of the models, the four-fold cross validation was applied. **Results:** For the training sets, the average mean absolute errors over three bulk densities of 1.30, 1.45 and 1.60 g cm⁻³ were 1.7578, 0.6157, 0.3812, 0.3113, 0.0103 and 0.0044% Wb for the linear regression, second-, third-, fourth-order polynomial regression, MLP and SVR models, respectively. For the validation sets, the average mean absolute errors over the three bulk densities were 1.7616, 0.6192, 0.3844, 0.3146, 0.0126 and 0.0093% Wb for the linear regression, 2nd, 3rd and 4th-order polynomial regression, MLP and SVR models, respectively. **Conclusion:** The regression models based on MLP and SVR yielded better performances than the models based on linear regression and polynomial regression on both training and validation sets. The models based on MLP and SVR also provided robustness to the variation of bulk density. Not only for dried longan aril, the proposed models can also be adapted and applied to other materials or dried food products.

Key words: Dried longan aril, dielectric constant, moisture content, bulk density, regression models, polynomial regression, linear regression, artificial neural network, multilayer perceptrons, support vector regression, cross validation, *Dimocarpus longan*

INTRODUCTION

Longan fruit (*Dimocarpus longan* Lour.) is a non-climacteric subtropical fruit grown commercially in many countries including China, Thailand, India, Vietnam, Australia and the United States (Jiang *et al.*, 2002). One of the factors that affect the deterioration and ultimately the costs of all dried agricultural products is moisture content (Karathanos, 1999). The moisture content of dried longan must not exceed 13.5% Wb (wet basis) of the whole fruit or 18% Wb of

aril according to the National Bureau of Agricultural Commodity and Food Standard of Thailand. Because no specific moisture content tester for dried longan was available, farmers would have to subjectively estimate the moisture from surface of skin, aril and seed by their experiences. Therefore, we previously proposed a prototype economical moisture measurement system for dried longan. It estimated the dielectric constant of dried longan aril by measuring the electrical capacitance of the dried longan-based capacitor. The system was designed to predict the moisture content of

dried longan aril from its dielectric constant by using the second-order polynomial.

There have been many research works in the area of system identification of a nonlinear black-box model. One of the most popular classes of artificial neural networks is the Multilayer Perceptrons (MLP) with the backpropagation algorithm as the training method. The MLP have been applied to several areas, for example, agriculture (Effendi *et al.*, 2010), medicine (Benamrane *et al.*, 2005; Isa *et al.*, 2007; Eiamkanitchat *et al.*, 2010), face recognition (Rizon *et al.*, 2006), electric power systems (Benslimane *et al.*, 2006).

The Support Vector Machine (SVM) is one of the most successful algorithms based on the statistical learning theory (Vapnik, 1999; Christiani and Shawe-Taylor, 2000). It was originally developed to solve classification problems but recently extended to the domain of regression problems known as the Support Vector Regression (SVR) (Gunn and Brown, 1999). One of the advantages of the SVM is that it has a few free parameters to adjust and solving for its optimal model parameters can be achieved using any standard quadratic programming algorithms. This can be done in a short time and there is no local minimum. It is a powerful technique for solving the nonlinear function approximation problems. Moreover, the Structural Risk Minimization (SRM) in learning SVM algorithm is more powerful than the Empirical Risk Minimization (ERM) in the MLP. It has been shown in several applications that both SVR and MLP provided better regression performance than the linear regression and polynomial regression, e.g., in flood prediction (Theera-Umpon *et al.*, 2008), electric load forecasting (Pahasa and Theera-Umpon, 2008; Abd, 2009), drug concentration estimation (Sumonphan *et al.*, 2008), power systems (Boonprasert *et al.*, 2003), computer networks (Hasegawa *et al.*, 2001), telecommunications (Suyaroj *et al.*, 2009), finance (Song *et al.*, 2010), environment (Mileva-Boshkoska and Stankovski, 2007).

In this study, we investigate the applications of the MLP and the SVR to another regression problem. The regression models try to predict the moisture content of dried longan aril from the dielectric constant. The accuracy of each model was evaluated by using the Mean Absolute Error (MAE). The comparison was conducted on several regression models including linear regression, polynomial regression (second-order, third-order, fourth-order polynomials), MLP and SVR.

MATERIALS AND METHODS

Dried longan preparation: The dried longan arils were prepared using the conventional drying process. Fresh longan fruits cv. Dew were dried at 70°C for 13 h and then at 75°C for 20 h. After that, the temperature was adjusted to 65°C for 15 h or until the moisture content was reduced to 10% Wb. However, in our experiments, 25 h after the beginning of the drying process, random samples were taken out every 2 h. They were further dried at 70°C under vacuum for about 8 h or until their weights were constants. This was according to the official methods and recommended practices of the Association of Official Analytical Chemists (Horwitz, 2005). Their actual moisture contents were calculated by:

$$\text{Moisture content (\% Wb)} = \frac{\text{Weight before drying} - \text{Weight after drying}}{\text{Weight before drying}} \times 100\% \quad (1)$$

The dielectric constant of each dried longan aril was measured by our previously proposed moisture measurement system at the room temperature. Dried longan aril was placed between 2 stainless steel discs inside a cylindrical plastic container. The weights of arils placed into the cylinder were varied from 9, 10 and 11 g (equivalent to the bulk densities of 1.30, 1.45 and 1.60 g cm⁻³ (five different moisture contents of 10, 14, 18, 22 and 25% Wb were considered. For each of the five moisture contents and each of the three bulk densities, 100 samples of dried longan aril were tested. Therefore, the total of 1500 samples were used in the experiments.

Artificial neural networks and support vector regression are well-described in literatures (Vapnik, 1999; Christiani and Shawe-Taylor, 2000; Haykin, 2008). We provide only their brief introduction in this study.

Artificial neural network: An Artificial Neural Network (ANN) is a mathematical model mimicking the biological neural network. An ANN can be considered as a universal function approximator and has been applied to several areas of research such as military, medicine, business. The typical structure of a feed forward neural network is displayed in Fig. 1. The goal is to find the best set of weights (*w*) so that the outputs *o_{j,n}* are as close to the desired outputs *d_{j,n}* as possible for a given input pattern *x_{i,n}*, *i* = 1, ..., *P* and *j* = 1, ..., *Q*. *P* and *Q* are the number of input features and the number of classes, respectively.

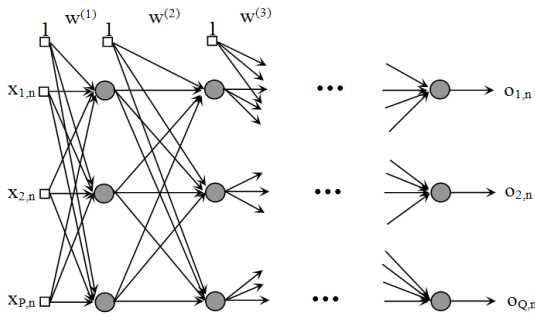


Fig. 1: Feed forward neural network

Support vector regression: In the support vector regression, the goal is to find a function $f(x)$ that has an ϵ -deviation from the actually obtained target y_i for all training set, $\{x_i, y_i\}$, $x_i \in \mathfrak{X}^n$, $y_i \in \mathfrak{R}$ with 1 observations. At the same time, $f(x)$ is as flat as possible. Suppose $f(x)$ takes the following form:

$$f(\mathbf{x}) = \langle \mathbf{w} \bullet \mathbf{x} \rangle + b \quad (2)$$

Where:

- w = A weight vector
- x = An input vector
- $\langle \bullet \rangle$ = The dot product
- b = A bias

Therefore, the objective is to choose a hyperplane that minimizes the Euclidean norm vector $\|w\|$ while simultaneously minimizes the sum of the distances from the data points to the hyperplane. By introducing $2n$ Lagrange multipliers α_i, α_i^* and using the Karush-Kuhn-Tucker (KKT) theory (Christiani and Shawe-Taylor, 2000). We obtain:

$$\mathbf{w} = \sum_{i=1}^1 (\alpha_i - \alpha_i^*) \mathbf{x}_i \quad (3)$$

Substituting Eq. 3 into 2 yields the regression function:

$$f(\mathbf{x}) = \sum_{i=1}^1 (\alpha_i - \alpha_i^*) \langle \mathbf{x}_i \bullet \mathbf{x} \rangle + b \quad (4)$$

where, x_i are the support vectors predetermined by the training patterns. From the KKT conditions, the support vectors are only the points x_i where exactly one of the Lagrange multipliers is greater than zero. For the nonlinear case, the input data need to be mapped into a high dimensional feature space. Let the nonlinear transformation function be $\Phi(\bullet)$ and using the kernel functions defined as:

$$K(\mathbf{x}_i, \mathbf{x}) = \langle \Phi(\mathbf{x}_i) \bullet \Phi(\mathbf{x}) \rangle \quad (5)$$

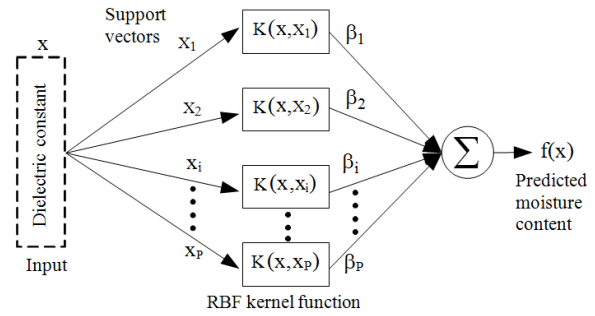


Fig. 2: Architecture of SVR-based moisture content prediction model

This implies that the dot product in the high dimensional space is equivalent to a kernel function of the input space. There are many types of kernel functions that can be used. The bias term b may be dropped if it is contained within a kernel function and the regression function in Eq. 4 is given by:

$$f(x) = \sum_{i=1}^1 (\alpha_i - \alpha_i^*) K(x_i, x) \quad (6)$$

$$f(x) = \sum_{i=1}^1 \beta_i K(x_i, x) \quad (7)$$

Moisture content prediction models: In this research, the input to the regression models is the dielectric constant of dried longan aril whereas the output is the moisture content. There are no parameters to set for the linear and polynomial regression models. However, there are some parameters to set for the MLP and SVR. For the MLP, back-propagation algorithm was applied in the training phase. Therefore, we needed to find the best structure for this particular problem, i.e., the number of hidden layers and the number of neurons in each hidden layer. For the SVR, the ϵ -insensitive loss function was applied. The support vectors are on the two hyper-planes with ϵ distance from the real hyper-plane. Therefore, ϵ is an error between an actual hyper-plane and the support vector hyper-planes. The data standing between the support vector hyper-planes are considered to produce no error. In the training stage, we try to find the support vector hyper-planes that can cover all training data. That is, all training data must be in between the two support vector hyper-planes.

Figure 2 shows the architecture of the prediction model. In this study, the Radial Basis Function (RBF):

$$K(\mathbf{x}_i, \mathbf{x}) = \exp\left(-\|\mathbf{x}_i - \mathbf{x}\|^2 / 2\sigma^2\right) \quad (8)$$

was used as the kernel function. Some important parameters to set for the SVR model include the RBF kernel specific parameter σ that controls the spread of the RBF and, therefore, the generalization of the SVR, the width ϵ of the tube and the regularization parameter C which controls the regression quality.

Evaluation procedures: To evaluate the system performance quantitatively, we use the Mean Absolute Error (MAE) which is defined as:

$$MAE = \frac{1}{n} \sum_{i=1}^n |Actual\ MC - Measured\ MC| \quad (9)$$

where, n is the number of the samples considered.

The experiments were conducted by using the four-fold cross validation technique which is a standard testing technique for any data set without training/test sets assignment. To be more specific, the entire data set (1500 samples) was randomly divided into four groups. Each group contains 375 samples, i.e., 75 samples for each of the 10, 14, 18, 22 and 25% Wb moisture contents.) In each validation, the data in each of four groups (called validation set) was used as the test set whereas the data in the three remaining groups were used as the training set. A regression model was generated by using the data in training set only. The derived model was then tested on the validation set to evaluate its generalization. Ultimately, after four validations, all samples in the data set are used as the test data in the validation sets. We evaluated the results using the average of MAE's over all four validations. The cross-validation was then performed on the data from each of the three bulk densities.

RESULTS

For parameter settings in the MLP and SVR, we performed extensive experiments to find the best sets of parameters under each condition. For the MLP, we found that two hidden layers with the numbers of hidden neurons of {3,5}, {5,9} and {4,6} yielded the best results for the bulk densities of 1.30, 1.45 and 1.60 $g\ cm^{-3}$, respectively, where the first and second elements of each pair denote the numbers of hidden neurons in the first and second hidden layers, respectively.

Furthermore, we found from many experiments that ϵ -insensitive loss functions with $\epsilon = 0.0001$ was the best choices for all three bulk densities of 1.30, 1.45 and 1.60 $g\ cm^{-3}$. The regularization parameter C , needed for solving the weight β_i , was chosen to be 100 for all three bulk densities. Finally, the parameter σ was set to 1.35, 0.65 and 0.85 for the bulk densities of 1.30, 1.45 and 1.60 $g\ cm^{-3}$, respectively.

The performances of the proposed models on the training sets of the four-fold cross validation are shown in Table 1. The average MAE's over the three bulk densities are 1.7578, 0.6157, 0.3812, 0.3113, 0.0103 and 0.0044% Wb for the linear regression, second-, third-, fourth-order polynomial regression, MLP and SVR models, respectively. Table 2 shows the performances on the validation sets. The average MAE's over the three bulk densities are 1.7616, 0.6192, 0.3844, 0.3146, 0.0126 and 0.0093% Wb for the linear regression, second-, third-, fourth-order polynomial regression, MLP and SVR models, respectively.

DISCUSSION

The results in Table 1 show that the polynomial regression models yield higher errors than the MLP and SVR models on the training sets of the four-fold cross validation by about one or two orders of magnitude. The MLP models also yield higher errors than the SVR models by about one order of magnitude.

The results in Table 2 are very similar to that on the training sets shown in Table 1. On the validation sets, the polynomial regression models yield higher errors than the MLP models by about one order of magnitude. In the mean time, the MLP models yield higher errors than the SVR models by about one order of magnitude.

It can be clearly seen that both SVR and MLP models yield better prediction performance than the models based on linear regression and polynomial regression in both training and validation sets. We can also see that the SVR models yield a little bit better performance than the MLP models in both training and validation sets. It is not surprising that the average MAE's of the training sets are less than that of the validation sets because the data in the validation sets are not involved in the model generation.

Table 1: Average mean absolute error of the training sets using four-fold cross validation

Bulk density ($g\ cm^{-3}$)	Average MAE (% Wb) (training sets)					
	Linear	2nd order	3rd order	4th order	MLP	SVR
1.30	1.5331	0.4302	0.3239	0.2499	0.0040	0.0030
1.45	1.8488	0.6703	0.3887	0.3341	0.0131	0.0016
1.60	1.8915	0.7466	0.4309	0.3500	0.0137	0.0086

Table 2: Average mean absolute error of the validation sets using four-fold cross validation

Bulk density (g cm ⁻³)	Average MAE (% Wb) (validation sets)					
	linear	2nd order	3rd order	4th order	MLP	SVR
1.30	1.5360	0.4334	0.3283	0.2530	0.0062	0.0043
1.45	1.8519	0.6725	0.3899	0.3356	0.0140	0.0100
1.60	1.8970	0.7517	0.4351	0.3551	0.0176	0.0136

Even though the average MAE's at the 1.30 g cm⁻³ bulk density are less than that at the other two bulk densities, the differences are not much in MLP and SVR models. Therefore, the system would have more robustness to the bulk density variation when the MLP or SVR is applied as the regression model.

CONCLUSION

The regression models based on multilayer perceptrons and support vector regression were proposed to predict the moisture content of dried longan aril from its dielectric constant. The performances of the proposed models were compared with that of linear regression and second-, third-, fourth-order polynomial regression models. The results using four-fold cross validation suggested that the SVR models achieved the best prediction performances, while the MLP models, polynomial regression models and linear regression models were next in line ordering from best to worst. The results also suggested that the bulk density of dried longan aril in the plastic container affected the prediction performances for the linear and polynomial regression models. However, this effect was very little when the MLP or SVR was applied. Therefore, both MLP and SVR models are the good choices for the system in that they provide very little prediction error and also provide robustness to the bulk density variation. Moreover, the proposed models are not only applicable to dried longan aril, they can also be adapted and applied to other materials or food products.

REFERENCES

Abd, M.K., 2009. Electricity load forecasting based on framelet neural network technique. *Am. J. Applied Sci.*, 6: 970-973. DOI: 10.3844/ajassp.2009.970.973

Benamrane, N., A. Freville and R. Nekkache, 2005. A hybrid fuzzy neural networks for the detection of tumors in medical images. *Am. J. Applied Sc.*, 2: 892-896. DOI: 10.3844/ajassp.2005.892.896

Benslimane, T., B. Chetate and R. Beguenane, 2006. Choice of input data type of artificial neural network to detect faults in alternative current systems. *Am. J. Applied Sc.*, 3: 1979-1983. DOI: 10.3844/ajassp.2006.1979.1983

Boonprasert, U., N. Theera-Umpon and C. Rakpenthai, 2003. Support vector regression based adaptive power system stabilizer. *Proceeding of the 2003 International Symposium on Circuits and Systems*, May 25-28, IEEE Xplore Press, USA., pp: 371-374. DOI: 10.1109/ISCAS.2003.1205033

Christiani, N. and J. Shawe-Taylor, 2000. *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. 1st Edn., Cambridge University Press, Cambridge, ISBN: 9780521780193, pp: 189.

Effendi, Z., R. Ramli and J.A. Ghani, 2010. A back propagation neural networks for grading *Jatropha curcas* fruits maturity. *Am. J. Applied Sci.*, 7: 390-394. DOI: 10.3844/ajassp.2010.390.394

Eiamkanitchat, N., N. Theera-Umpon and S. Auephanwiriyaikul, 2010. Colon tumor microarray classification using neural network with feature selection and rule-based classification. *Lecture Notes Elect. Eng.*, 67: 363-372. DOI: 10.1007/978-3-642-12990-2_41

Gunn, S.R. and M. Brown, 1999. SUPANOVA: a sparse, transparent modeling approach. *Proceeding of the 1999 IEEE Signal Processing Society Workshop on Neural Networks for Signal Processing*, Aug. 23-25, IEEE Xplore Press, Madison, WI., USA., pp: 21-30. DOI: 10.1109/NNSP.1999.788119

Hasegawa, M., G. Wu and M. Mizuni, 2001. Applications of nonlinear prediction methods to the internet traffic. *Proceeding of the IEEE International Symposium on Circuits and Systems*, may 6-9, IEEE Xplore Press, Sydney, NSW., Australia, pp: 169-172. DOI: 10.1109/ISCAS.2001.921273

Haykin, S., 2008. *Neural Networks and Learning Machines*. 3rd Edn., Prentice Hall, Upper Saddle River, New Jersey, ISBN: 0131471392, pp: 936.

Horwitz, W., 2005. *Official Methods of Analysis of AOAC International*. 18th Edn., AOAC Intl, Gaithersburg, Maryland, ISBN: 0935584803, pp: 34.

Isa, N.A.M., E. Subramaniam, M.Y. Mashor and N.H. Othman, 2007. Fine needle aspiration cytology evaluation for classifying breast cancer using artificial neural network. *Am. J. Applied Sci.*, 4: 999-1008. DOI: 10.3844/ajassp.2007.999.1008

- Jiang, Y., Z. Zhang, D.C. Joyce and S. Ketsad, 2002. Postharvest biology and handling of longan fruit (*Dimocarpus longan* Lour.) Postharvest Biol. Technol., 26: 241-252. DOI: 10.1016/S0925-5214(02)00047-9
- Karathanos, V.T., 1999. Determination of water content of dried fruits by drying kinetics. J. Food Eng., 39: 337-344. DOI: 10.1016/S0260-8774(98)00132-0
- Mileva-Boshkoska, B. and M. Stankovski, 2007. Prediction of missing data for ozone concentrations using support vector machines and radial basis neural networks. Informatica, 31: 425-430. http://www.informatica.si/PDF/31-4/08_Mileva-Prediction.pdf
- Pahasa, J. and N. Theera-Umpon, 2008. Cross-substation short term load forecasting using support vector machine. Proceeding of the 5th International Conference on Electrical Engineering/Electronics Computer, Telecommunications and Information Technology, May 14-17, IEEE Xplore Press, Krabi, pp: 953-956. DOI: 10.1109/ECTICON.2008.4600589
- Rizon, M., M.F. Hashim, P. Saad, S. Yaacob, 2006. Face recognition using eigenfaces and neural networks. Am. J. Applied Sci., 3: 1872-1875. DOI: 10.3844/ajassp.2006.1872.1875
- Song, X., Y. Ding, J. Huang and Y. Ge, 2010. Feature selection for support vector machine in financial crisis prediction: A case study in China. Exp. Syst., 27: 299-310. DOI: 10.1111/j.1468-0394.2010.00546.x
- Sumonphan, E., S. Auephanwiriyaikul and N. Theera-Umpon, 2008. Interpretation of Nevirapine concentration from immunochromatographic strip test using support vector regression. Proceeding of the IEEE International Conference on Mechatronics and Automation, Aug. 5-8, IEEE Xplore Press, Takamatsu, pp: 633-637. DOI: 10.1109/ICMA.2008.4798830
- Suyaroj, N., N. Theera-Umpon and S. Auephanwiriyaikul, 2009. A comparison of NN-based and SVR-based power prediction for mobile DS/CDMA systems. Proceeding of the International Symposium on Intelligent Signal Processing and Communications Systems, Feb. 8-11, IEEE Xplore Press, Bangkok, pp: 1-4. DOI: 10.1109/ISPACS.2009.4806718
- Theera-Umpon, N., S. Auephanwiriyaikul, S. Suteepohnwiroj, J. Pahasha and K. Wantanajittikul, 2008. River basin flood prediction using support vector machines. Proceeding of the IEEE World Congress on Computational Intelligence, June 1-8, IEEE Xplore Press, Hong Kong, pp: 3039-3043. DOI: 10.1109/IJCNN.2008.4634227
- Vapnik, V., 1999. The Nature of Statistical Learning Theory. 2nd Edn., Springer, New York, ISBN: 0387987800, pp: 314.